THE AMERICAN JOURNAL OF PSYCHOLOGY

4789

EDITED BY

KARL M. DALLENBACH
University of Texas

AND

M. E. BITTERMAN
Institute for Advanced Study

E. B. NEWMAN
Harvard University

WITH THE COÖPERATION OF

Vol. LXVIII
MEZES HALL, THE UNIVERSITY OF TEXAS
AUSTIN, TEXAS
1955

3871

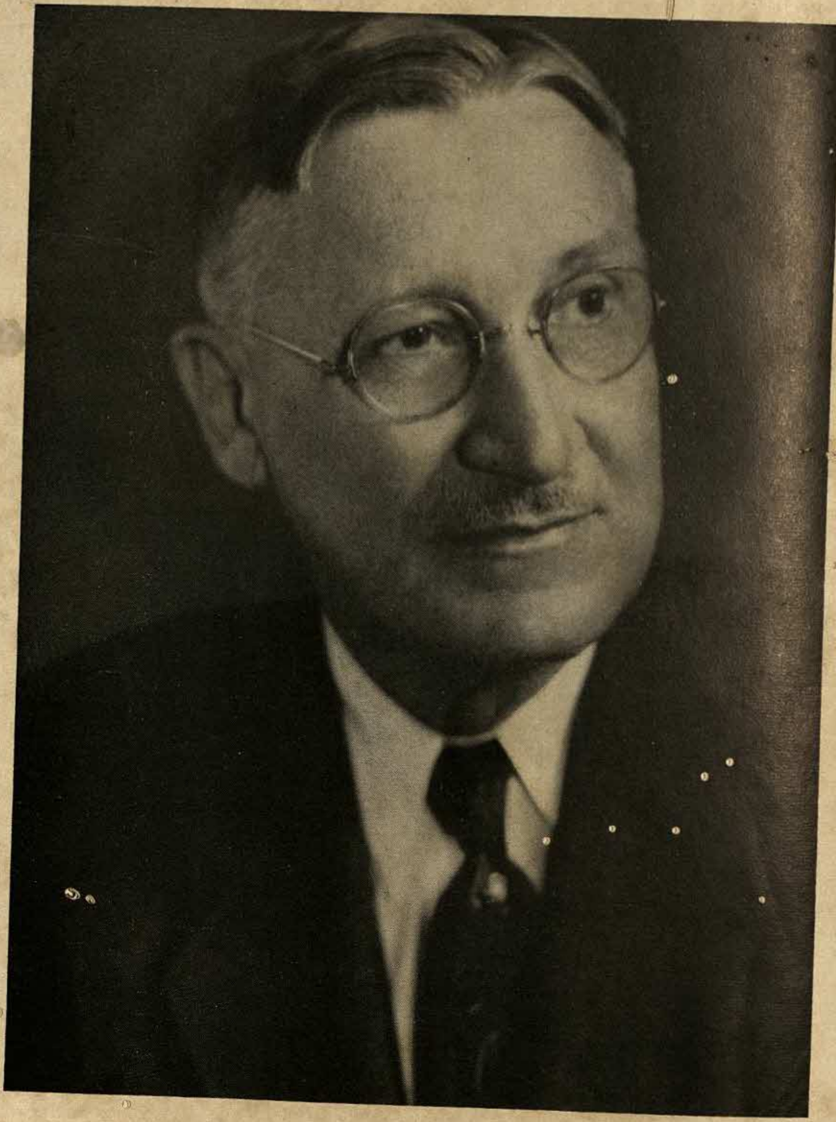
TABLE OF CONTENTS

Articles and Notes

BOOK REVIEWS

(The reviewer's name appears in parentheses after the title of the work.)

478g



Harvey A. Carr

(See page 149)

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. Stanley Hall

4788

Vol. LXVIII MARCH, 1955 No. 1

THE RELATION OF AREA AND LUMINANCE TO THE THRESHOLD FOR CRITICAL FLICKER FUSION

By Solomon Kugelmass and Carney Landis, Columbia University

Investigators have shown that many factors may influence the threshold for critical flicker fusion (*CFF*). Among these determinants are some that are primarily physical, such as the luminous flux coming from the test-patch; some that are primarily physiological, such as the O_2 level of the blood; and some that are psychological, such as practice. In some instances the detailed specifications of systematic relationships between more than one of these various factors and the *CFF*-threshold remains to be determined.

The interrelation of frequency, luminance, and area is of obvious importance in the determination of the *CFF*-threshold. The quantitative relationship between *CFF* and luminance or photometric brightness was first systematically described by Ferry[1] and by Porter.[2] Their findings first demonstrated that there is a linear relationship between *CFF* and log luminance. This relationship takes the form of an equation: $CFF = a \log I + b$, where $I =$ the luminance of the test-patch, while a and b are constants.[3] Although this relationship holds for medium luminances and frequency-values between 20 and 40 to 50 cycles per second, as higher luminance values are reached, the straight line tends to become negatively accelerated until a point is reached ($F_{max.}$) where further increase of luminance fails to increase the *CFF*-threshold, *i.e.* the test-patch continues to appear steady. The equation also

* Accepted for publication March 9, 1954. From the New York State Psychiatric Institute, Columbia University. This study was made possible by funds granted by the Carnegie Corporation of New York.

[1] E. S. Ferry, Persistence of vision, *Amer. J. Science,* 44, 1892, 192-207.

[2] T. C. Porter, Contributions to the study of flicker, II, *Proc. Roy. Soc.,* 70 A, 1902, 313-329.

[3] Throughout this article, $F =$ frequency in cycles per second (cps.); $I =$ luminance in millilamberts; and $A =$ area, either in square millimeters of retina surface or in visual angle subtended on retina.

fails to describe the data for values below about 20 cycles per second, except under certain special conditions.

In 1930, Granit and Harper described a similar straight line relationship between *CFF* and the logarithm of the area of the test-patch.[4] Working with several areas ranging between 1° and 5° of retinal visual angle they derived the following equation: $CFF = c \log A + d$, where A = area, and c and d are constants. They stated that "the complete curve relating fusion-frequency to log area is S-shaped,"[5] and that their equation was but an approximation. They further stated that this $CFF - \log A$ curve would tend to become horizontal at both the lower and upper ends. Piéron, using targets subtending visual angles ranging from 30″ to 0.5°, provided data which show the tendency toward the flattening of the $CFF - \log A$ curve at the smaller visual angles.[6]

Hecht and Smith, two observers in the same study, investigated the $CFF - \log I$ relationship utilizing four test-patches subtending visual angles of 0.3°, 2°, 6°, and 19°.[7] Some of their curves follow the $CFF = c \log A + d$ relationship but for the most part they show a wide range of inter-individual variation.

Allen varied the area of the stimulus-patch from approximately 10′ to about 6°.[8] His results when plotted did not constitute one straight line but rather a series of straight lines. The breaks between these lines he related to the change in the relationship of the distribution of the cones and rods in the retina.

We have assembled the findings of Granit and Harper, Piéron, Hecht and Smith, and Allen, together with one of our own curves and plotted them for $CFF - \log A$ in mm.[2] on the retina. The results are shown in Fig. 1.

The investigations cited above together with Fig. 1, indicate the following: (1) The relationship of *CFF* to log luminance is probably linear within certain limits. (2) The relationship between *CFF* and log area of a centrally fixated test-patch has been described as a straight line between approximately 1′ and 5° of visual angle. (3) It is probable, though not clear from available data, that the entire $CFF - \log A$ curve takes a sigmoid form and can be described as a negatively accelerated curve beyond 5° of visual angle. (4) Allen's data indicate that there are breaks in the $CFF - \log A$ curve which may be related to retinal histology.

Object. In the light of the foregoing we have studied more intensively the relationship between the area and luminance of a centrally fixated patch as they act to determine the threshold for *CFF*. By extending the range of areas somewhat beyond 14° of visual angle, and increasing the number of areas between 1° and 15° of retinal visual angle to be studied, it was hoped to obtain answers to the following questions.

---

[4] R. Granit, and P. Harper, Comparative studies on the peripheral and central retina: II. Synaptic reaction in the eye, *Amer. J. Physiol.*, 95, 1930, 211-228.
[5] *Ibid.*, 215.

[6] Henri Piéron, L'influence de la surface rétinenne en jeu dans une excitation lumineuse intermittente sur la valeur des fréquences critiques de papillotement, *Compt. R. Soc. Biol.*, 118, 1935, 25-28.

[7] Selig Hecht, and E. L. Smith, Intermittent stimulation by light: VI. Area and the relation between frequency and intensity, *J. Gen. Physiol.*, 19, 1936, 979-991.
[8] F. Allen, The delineation of retinal zones with dark tube vision, *Canadian J. Res.*, 23A, 1945, 21-31.

(1) Does the linear relationship between *CFF* and the logarithm of the area of the test-patch tend to fail, *i.e.* to become negatively accelerating as areas above 5° are studied?

(2) Are there definite breaks in this function which can be related to the inclusion of rods in the retinal area studied?

(3) What is the effect of the interaction of varying *F, I,* and *A* in differing systematic fashions? If *F, I,* and *A* are conceived as three dimensions in space, what is the appearance of the surface generated by the *CFF*-threshold?

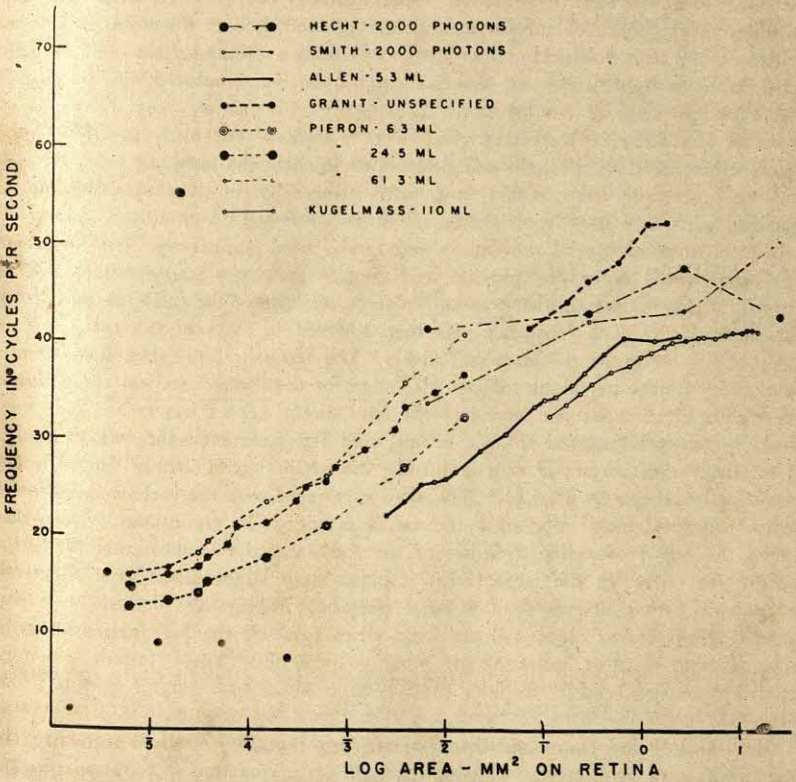


FIG. 1. THE *CFF* FOUND BY VARIOUS INVESTIGATORS AT DIFFERENT LEVELS OF LUMINANCE

Method. Because the Ferry-Porter law adequately describes the data between 20 and 40 cps., values of luminance and area were chosen which would include this range together with an extension at either end so that the total frequency range included was 14 to 46 cps. To determine the shape of the *CFF* — log *A* relationship beyond the 5° of visual angle which Granit and Harper employed, the

areas that were chosen ranged from 1.27° to 14.60° of visual angle. It was found that a luminance-range covering about 2.5 log units (1.74 to 435 ml.) permitted study of these areas within the chosen frequency-range.

If the frequency of intermittence is the dependent variable then one may select any one of the possible levels of luminance and vary the area (*A*) systematically or one may select any one of the possible areas and vary luminance (*I*) systematically, so determining the *CFF*-threshold in an orderly fashion. Data were obtained in both ways in different experimental series here reported. In both methods the sequence of presentation always proceeded from lower to higher luminances, and from smaller to larger areas so partially avoiding certain complications attributable to light-adaptation. When a single area was employed, the luminance was varied in steps of 0.2 log units. When the luminance was kept constant, 36 different size areas of increasing value were presented. For every area-luminance combination the *CFF*-threshold was determined in the same fashion.

Since proceeding from 'flicker' to 'steady' gives different thresholds from those obtained when one goes from 'steady' to 'flicker,' this second procedure, commonly called the method of the descending order, was used exclusively. The electronic apparatus which served to provide and control the light source was set at a point in cycles per second that was known to be well above the fusion-threshold for the particular *O*. The frequency was then lowered in steps of 0.5 cps. until *O* changed his report from 'steady' to 'flicker.' The frequency was then lowered one more step so that he might confirm his judgment of 'flicker,' and on confirmation the frequency in cycles per second of the first 'flicker' report was recorded. *E* then made the next luminance or area setting, and the same procedure was repeated. A relatively uniform set of criteria for the discrimination of 'steady' from 'flicker' was used by all our *O*s after each had some experience with the various appearances which different combination of *F*, *I*, and *A* provided. In preliminary tests it was found that onset of apparent flicker of the patch varied in appearance with areas of different sizes. In the larger areas it began as a slight glimmering about the edges, which would precede the usual pulsating appearance of flicker. It was agreed, therefore, to report 'flicker' only when most of the field appeared to be covered with a clear pulsating or moving brightness. These criteria provided relatively constant results through the whole range of areas and luminances employed.

Since Granit and Hammond demonstrated that the *CFF*-threshold is affected by the duration of observation,[9] finding that *CFF* is at maximal *F* at approximately 1-sec. duration of observation, we limited all single observations to 1 sec.

The experiment was carried out in a sound-deadened, air-conditioned room in which the level of illumination was approximately one foot-candle. Upon entering the experimental room, *O* seated himself in front of the apparatus, and light adapted to this level for about 5 min. *O* then rested his face on a 'hood' over the oculars of the microscope through which he made all his observations. He adjusted the interocular distance to provide binocular fixation. He then further light adapted

---

[9] R. Granit, and E. L. Hammond, Comparative studies on the peripheral and central retina: V. The sensation-time curve and the course of the fusion frequency of intermittent stimulation, *Amer. J. Physiol.*, 98, 1931, 654-663.

to the luminance of the surround provided by the apparatus (4.35 ml.) for about 5 min. before commencing the experiment.

Observers. Data were obtained from three practiced *O*s, *CL, SK,* and *ES,* who were 56, 26 and 25 yr. old, respectively. The observations were made at approximately the same time each day by each *O*. It was apparent from pilot studies and previous experimental data that with the range of values of luminance and area utilized, *SK* and *ES* would provide reasonably similar *CFF*-values. *CL,* on the other hand could be expected to be somewhat lower in his *CFF*-values with respect to both luminance and area.

Apparatus. The light source was provided by a glow modulator tube (Sylvania Type R1131C). This tube was powered by an electronic driver which caused it to deliver square pulses of light of carefully controlled duration and frequency of intermittence. A block diagram of the circuit appears in Fig. 2. The various circuit elements were supplied the proper voltages by electronically regulated

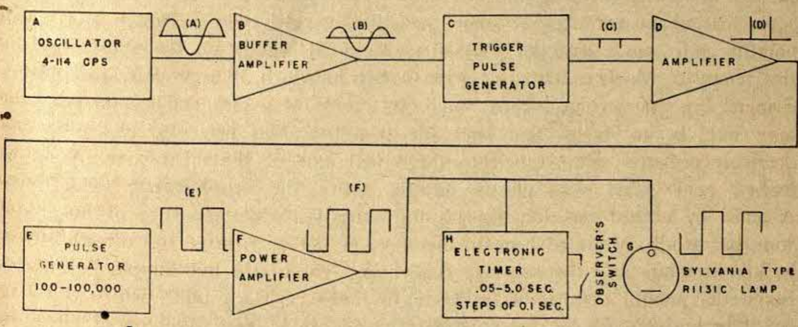


FIG. 2. BLOCK DIAGRAM OF THE APPARATUS

power supplies. When the proper voltages were applied to a Wien-bridge type oscillator at A, sine wave oscillations appeared at the output of the unit corresponding in frequency to the setting of the dials which were calibrated from 4 to 114 cps. in decade steps of 0.1 cps. If the vertical plates of an oscilloscope were attached to A, the wave-form illustrated (A) appeared on the screen. This wave was then passed through a buffer amplifier B, which provided isolation of the oscillator from the effects of loading by the succeeding circuitry. The trigger pulse generator followed, consisting of a neon lamp which ionized once each cycle of the wave-form at A. Differentiating the resulting wave-form produced the trigger pulse shown at (C), which was about 7 μ sec. wide. A second amplifier, D, inverted this wave and passed it with the wave form shown at (D), to the pulse generator, E. The pulse generator was a precision monostable multivibrator, producing the pulses shown at (E), which were adjustable in decade steps 10 μ sec. from 100 μ sec. to 100,000 μ sec. These then were the wave-forms of the desired variable characteristics. They were not, however, of sufficient power to activate the glow modulator tube. Therefore, they were applied to a power am-

plifier F, which supplied the required amplification. The resulting wave-form illustrated at (F), was applied to the glow modulator tube whose light output was linear with respect to the current supplied to it so that the resulting light output was that indicated to the right of G. To limit the duration of observation an electronic timer H, was so incorporated into the circuit that when *O* depressed his key, located on an extension-cable, a relay in the timer-circuit was activated which permitted the flash generator to operate until the presenting duration of 1 sec. elapsed, at which time the relay returned to its normal position causing the flash generator to cease operation. With this instrument, then, it was possible to deliver a carefully controlled light-pulse in which the duration and frequency of intermittence were accurately known and controlled. Although light–dark ratios adjustable from 0.0004 to 0.75 could be obtained, the light–dark ratio was maintained at 0.5 throughout the experiment by the proper settings of the dial controlling the duration of the flash and the frequency of repetition.

Two separate and complete calibrations of the instrument were carried out before the experimental series were undertaken. To insure continuously accurate performance, the driver was always monitored by an oscilloscope.

The optical system used to view the stimulus-patch was a Bausch and Lomb binocular microscope (Model TBV-8L). Mounted rigidly on the stage was the 'area template' which consisted of a brass strip in which 36 accurately spaced holes of increasing diameter had been bored. By means of a rack and pinion gear, the strip could be so moved that one hole at a time was presented and each was accurately centered with respect to the visual axis of the microscope. A bit of 'flashed opal' glass was placed directly above the area-template. This glass remained in a fixed position beneath the objective, while the strip of holes slid along beneath it. Mounted beneath the stage of the microscope and directly below the area-template was the substage condensing lens of the microscope. Below this lens was a neutral density wedge having a range of 3 log units, which could be moved back and forth by means of a rack and pinion gear. The top of the glow modulator tube was rigidly fixed below the neutral density wedge in such a way as to be centered with respect to the axis of the optical system. The beam of light emitted by the tube travelled through the neutral density wedge, the condensing lens, the hole in the area-template, and finally appeared on the opal glass as a circular patch of light of a selected luminance. The microscope was focused on this patch of light to a maximal definition for each *O*. The surround illumination was supplied by two 6-v. tungsten lamps operated from a storage battery and placed somewhat above the lower end of the microscope objective. The amperage (and hence the luminance) of these two lamps, connected in series, was maintained constant throughout the experiment by means of a rheostat and an ammeter.

O placed his head in a viewing hood so mounted that it could be comfortably adjusted to maintain the proper and constant distance between the eyes and the oculars of the microscope. Looking through the oculars, *O* would see a large circular field of light surrounded by darkness. In the center of this uniform light field was a smaller, darker circular patch. The size of this darker patch varied with the size of the hole in the area-template. It was this darker center which served as the test-patch when illuminated by the glow modulator tube. The largest test-patch employed measured 6.4 mm. at the surface of the flashed opal screen.

Its image subtended a visual angle of 14.6°. The visual angle was calculated taking into consideration the combined constant magnification of the objective and eyepieces (10x), and the projection distance which is 250 mm. The smallest hole measured 0.55 mm. at the flashed opal screen, and subtended 1.27° of visual angle. No artificial pupil was needed since the system of stops in the microscope so limited the cone of light that the image of the test-patch at the cornea was 2 mm. or less, which is considerably smaller than the natural pupillary diameter at the highest level of luminance employed in this study.

O fixated the center of the dark patch in the middle of the illuminated surround. When ready, he would press the button of the hand switch which he held. The dark area in the center would be illuminated for 1 sec. by the controlled intermittent light coming from the glow modulator tube below.

Luminance measurements of the test-patch and surround were obtained as follows. The neutral density-wedge was adjusted so as to place 1.5 log units of density between the glow modulator tube and the flashed opal glass screen. The glow modulator tube was then set to 60 cps. with each flash having a duration of 8330 μ sec. (*LDR* of 0.5) which provided an apparently steady light. *O* looking through the microscope varied the rheostat controlling the surround lamp-circuit until the surround matched the brightness of the test-patch. This value of the surround-luminance was then measured with a Macbeth Illuminometer. The average illuminometer measurement determined by 3 *O*s was 4.35 ml. This value of the surround was kept constant throughout the experiment, and different luminance-values of the test-patch were obtained by varying the calibrated neutral density-wedge. An experimental series during which area was systematically varied was run for each of five different luminance-levels: one below the level of the surround, one equal to the surround, and three higher than the surround.

Results. The relationships of *CFF* to log area in square millimeters on the retina at five different levels of luminance are presented in Table I and Fig. 3. Curves A, C, D, and E of Fig. 3 represent the averaged measures of *SK*, and *ES* at each area expressed as log area in square millimeters on the retina for log 2.64 ml., log 1.44 ml., log 0.64 ml., and log 0.24 ml., respectively. In these curves two determinations for each *O* (four measurements) were averaged between log 1.38 mm.^{2} and log 2.80 mm.^{2} and one observation on each *O* (two measures) between log 2.88 mm.^{2} and log 3.51 mm.^{2}

Curve B represents the average of six measures made on different days by *SK* with a luminance of log 2.04 ml. The shaded area around this curve (bounded by the smaller circles) contains the entire range of measurement obtained; it illustrates the slight amount of individual variability involved in these determinations.

The data for Curve B were subjected to the χ^2 test of linearity of regression to determine whether or not Curve B could be considered a straight line throughout its length. In other words does the Granit and

Harper law of linear relationship hold between log 1.38 mm.[2] and log 3.51 mm.[2] The results indicated that the regression is significantly non-

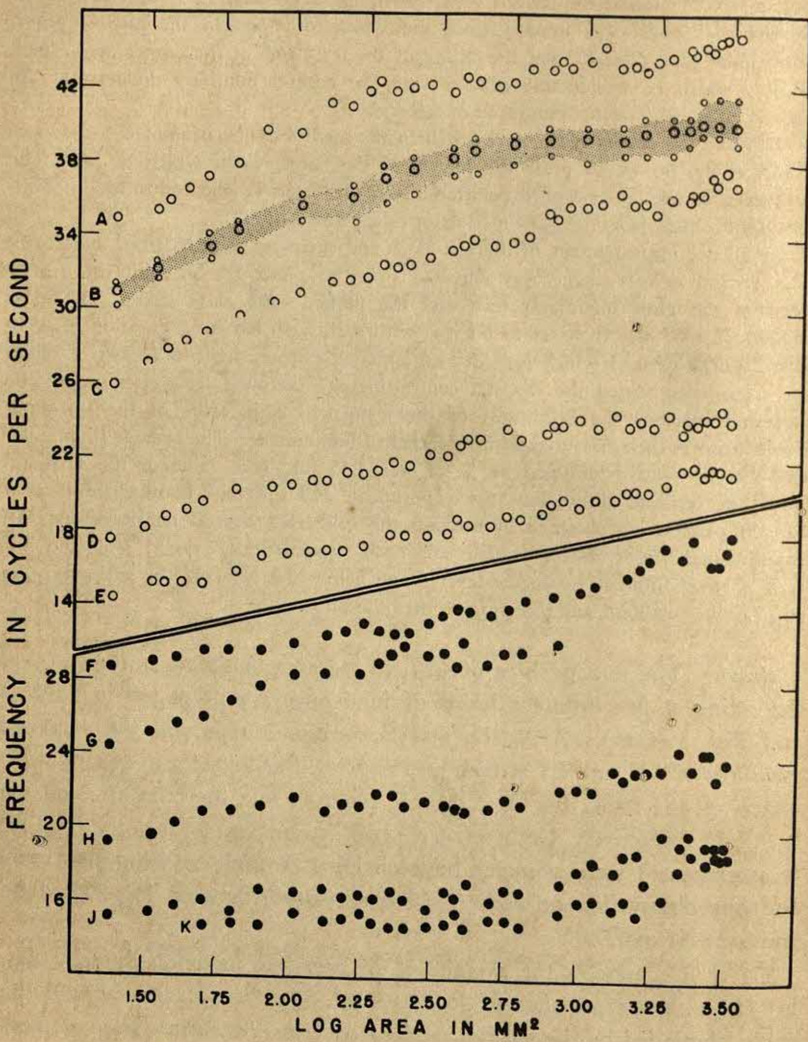


FIG. 3. RELATIONSHIP BETWEEN *CFF* AND LOG OF THE AREA IN SQUARE MILLIMETERS

linear at the 1% level. Employing the same test on the data below 2.76 log mm.[2] and the data above 2.76 log mm.[2], it was concluded that neither

of these two sections is significantly non-linear in form. It is thus possible to fit the data by two or more straight lines.

TABLE I

CFF THRESHOLDS FOR VARIOUS LUMINANCES (IN LOG MILLILAMBERTS) AND PATCH-AREA (IN DEGREES OF VISUAL ANGLE AND LOG AREA IN SQUARE MILLIMETERS)

(Columns A, C, D, and E give the averages of *ES* and *SK;* Column B, the average of six repetitions of *SK;* and Columns F, G, H, J, and K, the results of our oldest *O, CL.*)

Test-patches		Average *CFF* thresholds for various luminances									
		for *ES* and *SK*					for *CL*				
Visual angle (degrees)	log area (mm²)	2.64 A	2.04 B	1.44 C	0.64 D	0.24 E	2.64 F	2.04 G	1.44 H	0.64 J	0.24 K
1.27	1.38	35.0	32.3	26.0	17.5	14.5	28.8	24.5	19.3	15.3	—
1.50	1.52	35.5	33.5	27.5	18.8	15.4	—	—	19.8	15.5	—
1.83	1.70	37.4	34.6	28.8	19.8	15.5	29.8	26.3	21.0	16.3	—
2.07	1.80	38.0	35.6	29.8	20.4	16.0	—	—	21.0	15.5	—
2.30	1.90	39.7	—	30.3	20.5	16.9	29.8	27.8	21.3	16.8	14.8
2.63	2.02	39.8	36.8	31.0	20.8	17.0	—	—	21.8	16.5	—
2.98	2.12	41.2	—	31.6	21.0	17.1	30.5	28.5	21.3	16.8	15.0
3.20	2.19	41.0	37.4	31.9	21.4	17.1	—	—	21.5	16.5	15.3
3.43	2.25	41.9	—	31.9	21.4	17.4	31.3	28.5	21.3	15.5	15.0
3.67	2.30	42.5	38.4	32.6	21.5	18.0	—	—	22.0	16.3	15.0
3.90	2.36	42.0	—	32.5	22.0	18.0	30.5	29.5	22.0	16.8	14.8
4.13	2.41	42.3	38.9	32.6	21.8	18.0	—	—	21.3	16.3	14.8
4.47	2.48	42.3	—	33.0	22.5	18.0	31.3	29.5	21.8	15.8	14.8
4.80	2.54	42.0	39.4	33.4	22.4	18.1	—	—	21.5	16.8	15.0
5.03	2.58	42.8	—	33.8	22.9	18.9	32.0	28.5	21.3	16.3	15.5
5.27	2.62	42.5	39.9	33.9	23.4	18.6	—	—	21.3	17.3	14.8
5.73	2.69	42.4	—	33.8	23.5	18.5	31.8	28.8	21.3	16.3	15.3
6.17	2.76	42.5	40.1	33.9	23.9	19.3	—	—	21.8	17.0	15.3
6.53	2.80	43.2	—	34.1	23.4	19.0	32.5	29.8	21.5	16.8	15.0
7.10	2.88	43.3	40.4	35.5	23.8	19.3	—	—	—	—	—
7.33	2.91	43.7	—	35.3	24.0	19.8	—	—	—	—	—
7.57	2.93	43.3	40.4	35.8	24.3	20.0	33.0	30.0	22.3	17.3	16.5
8.23	3.01	43.8	—	35.8	24.5	19.8	34.5	—	22.5	18.0	16.5
8.70	3.05	44.5	40.3	30.0	24.0	20.0	34.5	—	22.5	18.5	16.5
9.37	3.12	43.3	—	36.5	24.8	20.0	34.5	—	23.5	18.0	16.0
9.83	3.16	43.5	40.7	36.0	24.0	20.5	34.5	—	23.0	19.0	16.5
10.30	3.20	43.3	—	36.0	24.3	20.5	34.5	—	23.5	19.0	15.5
10.78	3.24	43.8	40.8	35.5	24.0	20.5	33.5	—	22.5	19.0	18.0
11.43	3.29	40.0	—	36.3	24.3	21.0	35.5	—	23.5	20.0	16.5
12.10	3.34	44.5	40.9	36.3	23.8	21.8	36.5	—	25.0	19.0	19.0
12.33	3.36	44.0	—	36.5	24.3	21.5	35.0	—	24.5	19.5	18.0
12.77	3.39	44.5	41.2	36.5	24.3	22.0	35.5	—	23.5	20.0	19.0
13.23	3.42	44.3	—	37.5	24.5	21.5	34.5	—	24.5	20.0	17.0
13.70	3.45	44.8	41.2	37.0	24.5	21.8	34.5	—	24.5	19.5	18.5
14.13	3.48	45.0	—	37.8	25.0	21.8	34.5	—	23.0	19.5	19.0
14.60	3.51	45.0	41.0	37.0	24.3	21.5	36.0	—	24.0	19.5	19.0

Curves F, G, H, J, and K of Fig. 3 represent the measures obtained by the older *O, CL,* for luminances of log 2.64 ml., log 2.04 ml., log

1.44 ml., log 0.64 ml., and log 0.24 ml., respectively. (A sufficient number of determinations were not made in the establishment of the Curves A, C, D, E, F, G, H, and J to permit an adequate mathematical test for linearity.)

Examination of the material presented in Table I and Fig. 3 justifies the following statements.

(1) The $F - \log A$ curves for a single *O* represent a relatively stable function so far as day-to-day measures are concerned.

(2) The linearity of area function assumed by the Granit and Harper equation can be demonstrated in some *O*s at some levels of luminance between approximately 1° and 5° of visual angle.

(3) The $F - \log A$ curve is probably better described as a negatively accelerating function without significant or regular breaks between 1° and 15° of visual angle. That is, the Granit and Harper equation which implies linearity has real limitations.

(4) Granit and Harper found that the slope of the $CFF - \log A$ function was increased if the level of luminance was raised. This systematic relationship may also be noted in the present study (see Fig. 3). Although not conclusive, the data of the present study indicate the possibility that the $CFF - \log A$ curve tends to become horizontal (asymptotic) sooner with increased luminance. Since contributions to the total energy of the stimulus come from both the luminance-level, and the area of the test-patch, it is possible that at higher luminances, further increases of energy obtained by increasing the area of the test-patch cannot be as efficiently used by the visual mechanism.

(5) The data of *CL* (age 56 yr.) shown in Fig. 3, and of Hecht (age *ca.* 42 yr.) and of Allen (age *ca.* 65 yr.) shown in Fig. 1, indicate that increasing chronological age does not have a regular effect on the slope or the apparent curvilinearity of the $CFF - \log A$ function. The breaks found by Allen in his own observations are not duplicated as to place of break by the data which we obtained nor by the data of the other investigators as shown in Fig. 1. It is possible that breaks in functional relationship occur but that they are related to the incidence of cones and rods in the retina, such relationship will have to be determined by a more complete study directed toward this particular consideration.

The relationship of *CFF* to the logarithm of luminance for five different size patches (14.6°, 7.57°, 2.41°, 2.07° and 1.27° or log 3.51, 2.73, 2.41, 1.80 and 1.38 mm.^{2}) are shown in Fig. 4. Each point in the curves of *CL* represents one determination, while each point on the *AV* curves (the

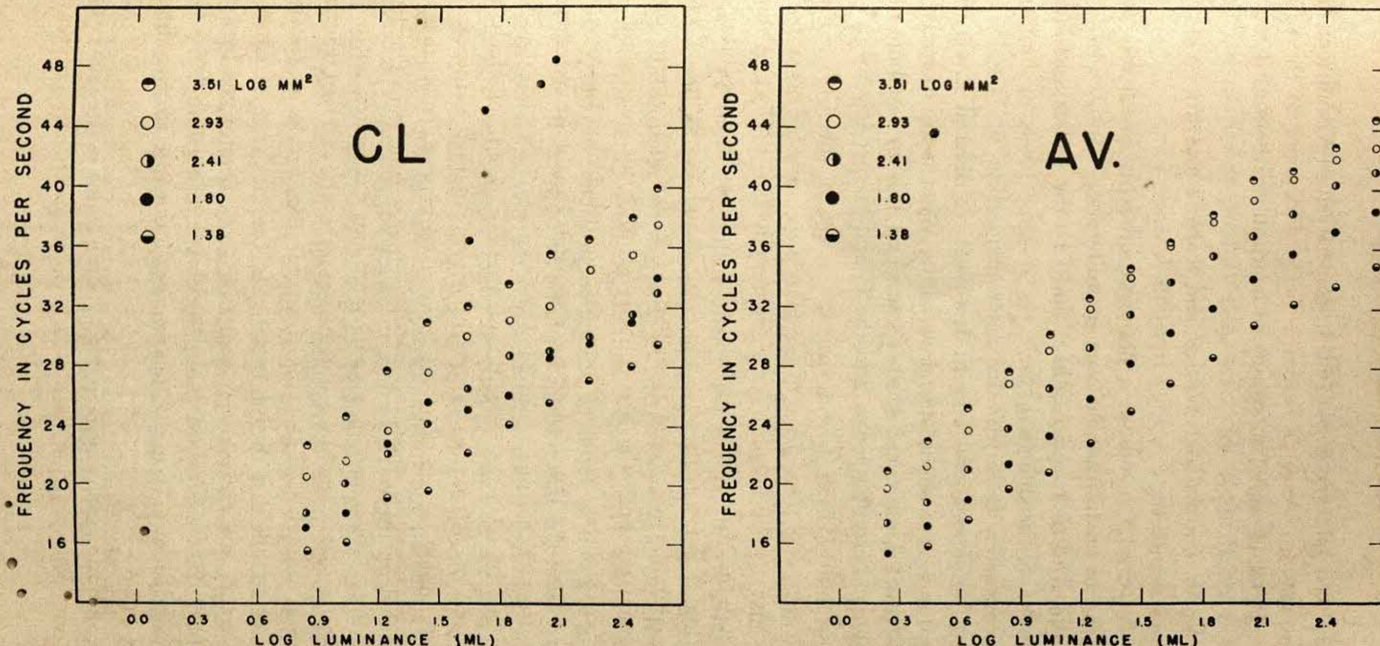


FIG. 4. RELATIONSHIP BETWEEN *CFF* AND LUMINANCE

The graph to the left gives the results for *CL;* the one to the right, the average results of *ES* and *SK.*

average of *ES* and *SK*) represents four trials, two for each *O*. Fig. 4 indicates the following.

(1) The Ferry-Porter law of linearity of function of $F - \log I$ holds for areas between visual angles of 1.27° and 14.60°; this is in line with Hecht and Smith, Crozier and Wolf,[10] and many other investigators of the influence of area on the $F - \log I$ function.

(2) The curves of *CL* indicate a decreased efficiency in the $F - \log I$ function. This sort of decrease has been found by McFarland, Holway and Hurvick,[11] and by Erlick and Landis,[12] and in each instance it was attributed to increased chronological age.

Thus far we have dealt with the interrelationships of frequency of intermittence, luminance, and area of test-patch considered in terms of any combination of two of those three variables. What is the interrelationship of all three? Granit and Harper expressed this by combining the equations of the Ferry-Porter law and the Granit-Harper law to give the following equation:

$$CFF = \alpha \log I \log A + \beta \log I + \gamma \log A + \delta$$

where I = intensity in candles per square meter, A = area in mm.² of retinal image and α, β, γ and δ are constants which may be derived from the individual curves. We have tried to represent the relationship of *F, I,* and *A* geometrically. Figs. 5 B, 5 C and 5 D are three dimensional graphic representations drawn in terms of scaling in projective geometry for the data obtained from the observations made by *SK*. Figs. 5 B, 5 C and 5 D represent the 'surface' which resulted when the measurements were 'cut through' a cube whose three directional edges were scaled, right to left on the base in terms of log mm.² of retinal *area* stimulated, left to right on the base scaled in log millilamberts *luminance* of the test-patch and *frequency* scaled on the vertical edges in cycles per second. The resultant surface represents the interaction of *F, I* and *A* which give the *CFF*-thresholds. That one might have some idea of what this *F-I-A* surface might look like when measures were averaged and smoothed we utilized

---

[10] W. J. Crozier, and Ernst Wolf, Theory and measurement of visual mechanisms: IV. Critical intensities for visual flicker. Monocular and binocular, *J. Gen. Physiol.*, 24, 1941, 505-534.

[11] R. A. McFarland, A. H. Holway, and L. M. Hurvich, *Studies of Visual Fatigue*, 1942, 1-255.

[12] Dwight Erlick, and Carney Landis, The effect of intensity, light-dark ratio, and age on the flicker-fusion threshold, this JOURNAL, 65, 1952, 375-388.

the data on which Figs. 5 C and 5 D were based to derive the 'idealized' surface shown in Fig. 5 B.

A consideration of Fig. 5 B indicates: (a) that the slope of the *F-I-A* surface is pitched at a greater angle by *I* than by *A*. (b) Although as

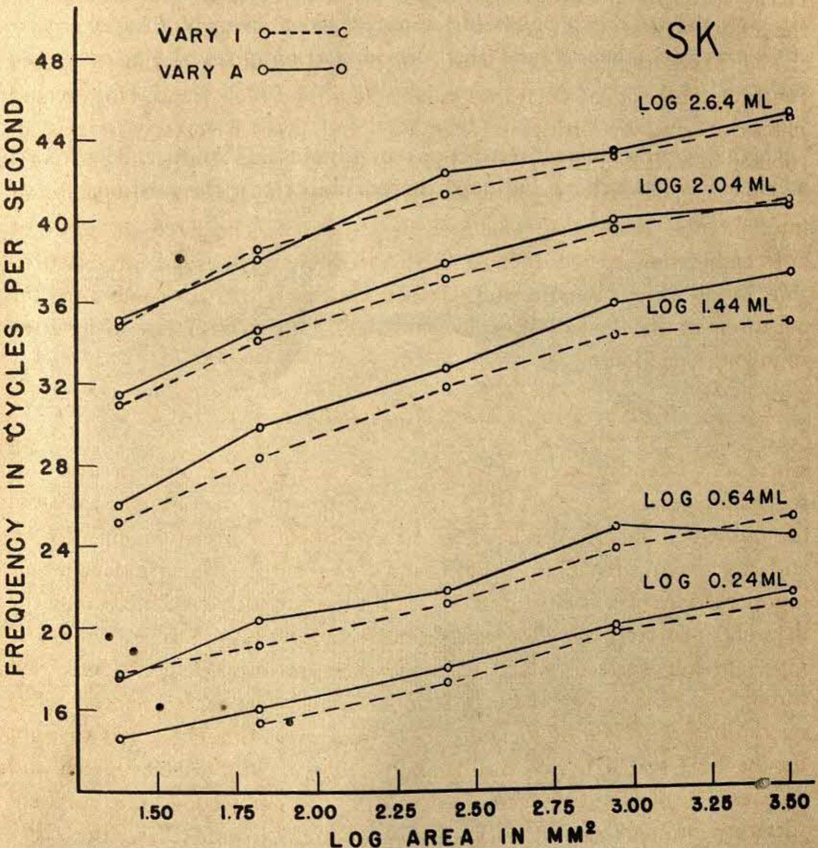


FIG. 5 A. GRAPHIC REPRESENTATION OF OBSERVATION OF *SK*

shown in Fig. 5 A, the Ferry-Porter equation and the Granit-Harper equation both seem to indicate that $F - \log I$ and $F - \log A$ are within limits straight line functions, the interaction of the three variables gives a surface which is curved with respect to each of the three variables *F, I,* and *A*.

We have made similar three dimensional graphs of the data which

we obtained from the other two *O*s, *CL* and *ES*. We have also made similar three dimensional graphs for *SH* and *ELS* from the data provided by Hecht and Smith and for *WJC* from assembled data provided by various publications of Crozier and Wolf.[13] These additional graphs although differing markedly among themselves and from that shown in Fig. 5 B all show the same two points just commented on; namely, *I* has a greater effect on slope than *A* and that the interaction gives rise to a curved surface.

Discussion. Previous investigations of Granit and Harper, Hecht and Smith, and Piéron, have indicated the possibility that the relationship be-

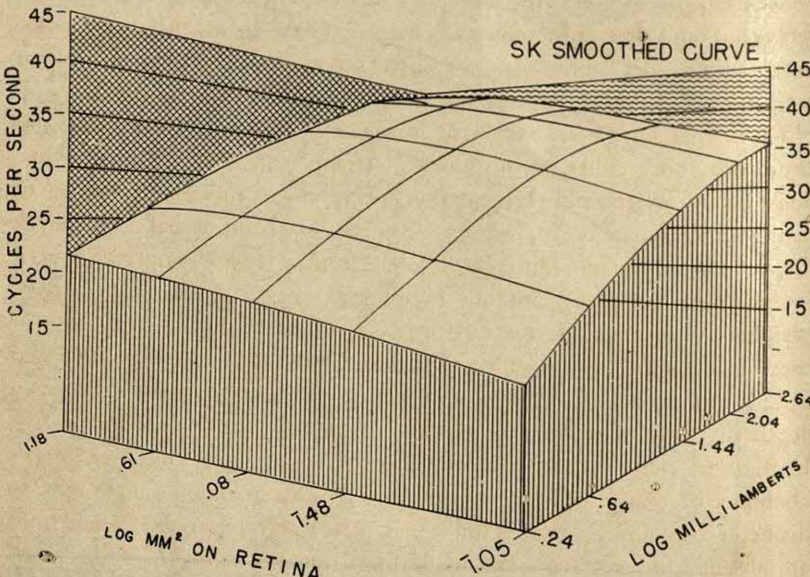


FIG. 5 B. TRI-DIMENSIONAL REPRESENTATION OF SMOOTHED DATA SHOWN IN FIGS. 5 A, 5 C, and 5 D.

tween $F - \log A$ is linear between 5′ and 5° of visual angle. The essential point of investigations of this sort is that, within the middle range of illuminance and of area, the *CFF*-threshold increases or decreases in arithmetic steps while objective radiance of luminous energy from the

[13] Carney Landis, Determinants of critical flicker fusion threshold. *Physiol. Rev.*, 34, 1954, 259-286.

test-patch and the size of the area about the visual axis of regard both increase or decrease in logarithmic steps. In other words the Ferry-Porter equation and the Granit-Harper equation are special cases of the Weber-Fechner equation.

Our findings (Figs. 3 and 4) are in accord with those of previous investigators in that they indicate that the effect of changing area within the range of 1.27° and 14.60° does not alter the Ferry-Porter equation in the range of frequencies and luminances utilized. Granit and Harper did not state the $F - \log A$ relationship as a 'law' and they did indicate that it had limited applicability. That the $F - \log A$ relationship is approximately linear for some individuals and over an approximate range between 1° and 5° seems apparent just as the $F - \log I$ relationship is approximately linear between 1 and 100 millilamberts. There is, however, a fairly wide range in the limits of applicability of these equations in terms of area or luminance, and a wide range of variation among individual *O*s.

The retinal surface on which the image of the test-field is formed is not a uniform one, either in histological arrangement of visual receptors, thickness of overlaying pigmentation, nerve fibers and blood vessels, or in the angular relation of the sense receptors with the visual axis. Polyak divides the retina into different regions; central fovea, parafoveal, perifoveal, near periphery, middle periphery, far periphery, and extreme periphery.[14] In man, the central fovea is a small pit-shaped depression or excavation in the inner face of the retina in the direct line of the visual axis. It corresponds to a visual angle of 5° (or slightly more) and has a depth of 240μ. The floor of the pit (foveola) corresponds to 1° 20″. In the very center of the foveola the cones become very thin and long, a "bouquet of central cones"[15] constituting an area of 0.005-0.0075 mm. in diameter containing about 2500 cones. The central territory where rods are absent measures about 0.5 to 0.6 mm. across (1° 40′ to 2° of visual angle) containing approximately 34000 cones which are arranged in a regular and uniform fashion. "The transition from one territory to the next in the central fovea and in its vicinity is usually gradual and barely noticeable, being abrupt only in one place, viz., where the foveal slope becomes its floor [1° 20′]. Nor is the inner fovea itself [5° of visual angle] marked off well enough by any clear boundary from the surrounding perifoveal region."[16] In brief there is a bouquet of central cones subtending a visual angle of 20′ to 30′. This is surrounded by a further rod

[14] S. L. Polyak, *The Retina,* 1941, 196 ff.
[15] *Ibid.,* 204. [16] *Ibid.,* 209.

free, uniform cone area, constituting in all a visual angle of 1° 40′ to 2°. This in turn is surrounded by a predominantly cone area extending to a total visual angle of 5° (or a little more). The perifovea corresponds to 18° 20′ and the parafovea to 24° to 26° of visual angle and contains both cones and rods. Polyak concludes, "In the final analysis it is, to a certain extent, a matter of preference what one chooses to declare to be the central fovea or any other territory in the retina. In any case, the connections or synaptical relationships of the superposed successive layers in the particular territories, since they point toward functional relationships, may serve as a more reliable criterion than some other more conspicuous but less essential feature."[17] Accepting Polyak as the outstanding authority on the anatomy of the retina his suggestions from the anatomical point of view would incline one to expect that if changes in the frequency-area function do occur they should be found at either or both, the region of visual angular measurements of about 10′-15′ (central bouquet) and at about 5°, which roughly correspond to log area mm² (Fig. 1) of log $\bar{3}.0$ and log 0.2.

Our investigation had no thresholds below 1° 20′. An examination of all the curves given in Fig. 1 shows that a break *may* occur near log $\bar{4}.0$ and $\bar{2}.0$ (data of Piéron and of Allen). Our investigations indicate that under some combinations of circumstances a break occurs at 5° or 6° (just beyond log 0, Fig. 1; log 2.5, Fig. 3; and log 0.08 or log 0.15 on Figs. 5 C and 5 D). Other curves shown in Fig. 1 show that such a break at log 0 may have been found by Allen, Granit and Harper, and Hecht and Smith, but not enough measures in this vicinity have been made on enough *Os* to make this more than a possibility. In our opinion the evidence indicates that in the midrange of luminance the *CFF*-log area relationship may be considered linear from about 15′ to about 5° of visual angle. Below 15′ the function is probably a positively accelerating curve and above 5° it is probably a negatively accelerating curve.

As a practical point many investigators have used test-patches of 30′—2° so as to insure 'pure cone functions.' Our analysis of this situation indicates that a reasonably straight line function exists between 15′ and 5°. If rods do enter histologically between 1° 20″ and 5° of retinal visual angle, they do not modify the *F-I-A* relationship under any circumstances so far investigated.

The studies of Hecht and Smith and of Allen may be interpreted as indicating that with increasing age increasing area has less effect in in-

[17] *Loc. cit.*

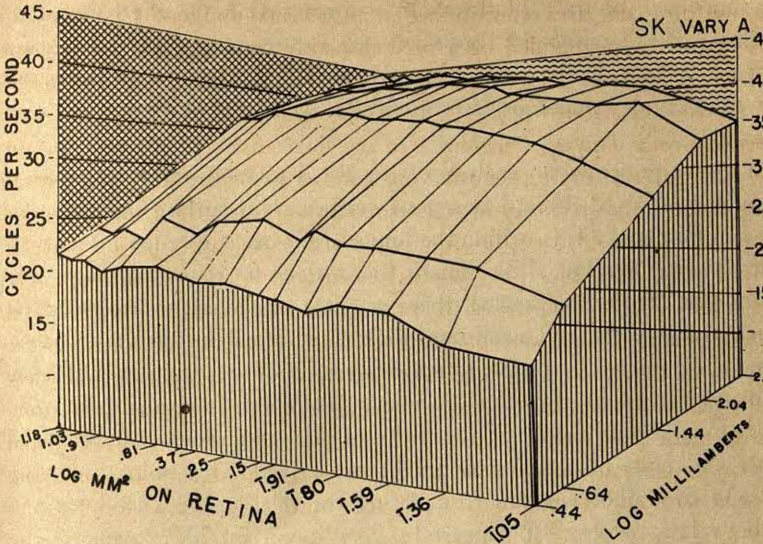


FIG. 5 C. TRI-DIMENSIONAL REPRESENTATION OF *SK*'S OBSERVATIONS WHEN AREA WAS VARIED

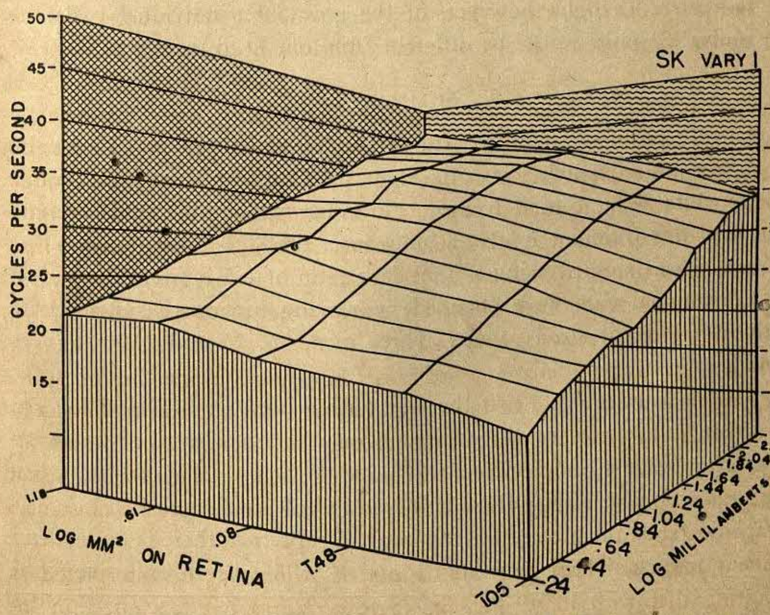


FIG. 5 D. TRI-DIMENSIONAL REPRESENTATION OF *SK*'S OBSERVATIONS WHEN LUMINANCE WAS VARIED

creasing the *F* of the *CFF*-threshold at a constant *I*. On the basis of all available data it cannot be concluded that increasing age, per se, has any systematic effect on the *CFF* – log *A* function. It is more probable that the effect is dependent on decreased effectiveness of *I* than it is on *F* or on *A*.

We have presented an analysis of the *F-I-A* relationship in terms of an empirically derived three dimensional surface. This surface indicates that log *I* in millilamberts is, within the limits so far investigated, a more active determinant than log *A* in square millimeters of retinal surface. This means that, as here expressed, there is not a one-to-one relationship between millilamberts and square millimeters of retinal surface. It also indicates that it might be possible to derive a system of area units in which one such unit would be equivalent to one millilambert. It also indicates that a one-to-one equivalence of the units of flash duration, units of luminance and units of retinal area might be derived. This point was made in 1834 by Talbot who, in formulating the relationship now known as Talbot's law, wrote: "It suggests a very important idea; namely, that *time* may be employed to measure the *intensity* of light."[18] We would add the possibility that *time* might also be employed to measure the *area* of the test-patch. In light, however, of the great inter-individual variability the equivalent units would be different from one *O* to the next.

SUMMARY

The relationship of *CFF* to area of the test-patch was examined using 36 different sized patches ranging from 1.27° to 14.60° of visual angle. Using white light, central fixation, binocular regard, surround illumination, and five different levels of luminance, the *CFF* –log *A* (area) relationship was studied using a light-dark ratio of 0.5 between 14 and 48 cps. Additional data were obtained by varying luminance systematically with five different sized patches. Three men, 56, 26 and 25 yr. of age, served as *O*s.

(1) The data indicate that the relationship between *CFF* and log area in square millimeters of the test-patch may best be described as an approximate straight line between 1.27° and 5° of retinal visual angle and as a negatively accelerating function beyond 6° of visual angle.

(2) So far as the area of the test-patch is concerned there is no evidence that any particular area between 15′ and 5° will give an exclusive 'cone'

[18] H. F. Talbot, Experiments on light, *Phil. Mag.*, 5, 1834, 321-334, esp. 328-329.

function. The action of rods, within the range of 1 and 100 millilamberts of white light does not show itself until past 5° of visual angle, if negative acceleration of the $F -$ log A curve is the criterion used.

(3) Our data do not clearly indicate any systematic discontinuities in the $CFF -$ log A function that can clearly be asociated with retinal histology. That such breaks or discontinuities in function occur is possible but insufficient evidence is available to provide positive assurance of the point.

(4) The geometric tri-dimensional surface of F in cycles per second, *log I* in millilamberts, and log A in square millimeters of retinal surface is curved. It is possible that one might establish an individual one-to-one relationship of units of time, luminance, and retinal surface stimulated at least so far as the *CFF*-threshold is concerned. If such a relationship were established the surface would be flat.

AN ATTEMPT TO DIFFERENTIATE BETWEEN INDIVIDUALS WITH HIGH AND LOW REASONING ABILITY

By OLGA W. MCNEMAR, Stanford University

This paper presents the results of an attempt to differentiate, within a college sample, individuals high in reasoning ability from those low in this ability. The variables to be studied may be classified under three general headings: (1) free and controlled word-association; (2) deduction and induction; and (3) experimentally induced set.

Although the writer's interest was primarily in problem-solving, the lack of appropriate tests of known factorial content and proven reliability made it advisable to take as the criterional variable a more readily measurable ability; namely, *logical reasoning*. This factor, which was recently isolated and named by Guilford, Green, and Christensen, seems to us an indispensable, if not a fundamental, aspect of problem-solving,[1] Guilford and his collaborators regard logical reasoning as identifiable with the factor of deduction, found earlier by Thurstone,[2] Blakey,[3] and Zimmerman,[4] but they rejected the name 'deduction' because tests saturated with this factor do not require the testee to make his own deductions, he had merely to choose from those given. Two subsequent studies by Guilford and his collaborators—a second analysis of reasoning abilities[5] and an analysis of evaluative abilities[6]—have confirmed the factor of logical reasoning. In all three investigations this factor was shown to involve "the sensitivity to logical relationships in the testing of the correctness of a conclusion."[7]

* Accepted for publication December 14, 1953. This study is a part of a series being carried out under Project 150-149 and supported by Contract N6onr 25125 between Stanford University and the Office of Naval Research. Work under the contract is under the general direction of Dr. Donald W. Taylor.

[1] J. P. Guilford, R. F. Green, and P. R. Christensen, A factor-analytic study of reasoning abilities: II. Administration of tests and analysis of results, *Rep. Psychol. Lab., Univ. S. Calif.*, 1951, (no. 3), 1-28.

[2] L. L. Thurstone, Primary mental abilities, *Psychomet. Monog.*, 1938, (no. 4), 1-121.

[3] R. I. Blakey, A factor analysis of non-verbal reasoning tests, *Educ. & Psychol. Meas.*, 1, 1941, 187-198.

[4] W. S. Zimmerman, A revised orthogonal rotational solution for Thurstone's primary mental abilities test battery, *Psychometrika*, 18, 1953, 77-94.

[5] J. P. Guilford, R. F. Green, P. R. Christenson, A. F. Hertzka, and N. W. Kettner, A factor analytic study of Navy reasoning tests with the Airforce Aircrew Classification Battery, *Rep. Psychol. Lab., Univ. S. Calif.*, 1952, (no. 6), 1-23.

[6] Guilford, Hertzka, and Christensen, A factor-analytic study of evaluative abilities: II. Administration of tests and analysis of results, *ibid.*, 1953, (no. 9), 1-28.

[7] *Ibid.*, 17.

Tests and Subjects

The criterion. Members of the research staff at Stanford who had been working for some time in the area of problem-solving were asked to examine the 34 tests of Guilford's first reasoning battery[8] and to check those which required the type and level of reasoning involved in problem-solving.

Eleven tests were checked by five of the seven judges, none of whom was familiar with the published factorial content of the tests. First centroid loadings, redetermined on the basis of the intercorrelations among the 11 selected tests, ranged from 0.43 to 0.69 with an average of 0.54. In the analysis by Guilford, Green, and Christensen these 11 tests had in general yielded larger projections on logical reasoning than on any of the other reasoning factors.[9]

Since it was desired that our test be as pure a measure as possible, the final rotated matrix reported by these authors was carefully scrutinized with the aim of selecting tests having high variance on the logical reasoning factor and low variances on verbal, numerical, and perceptual factors. One additional consideration, administrative time, necessitated limiting the number of tests to four.

The resulting criterional measure (hereafter to be referred to simply as the Reasoning Test) contains the following tests, described here briefly in the order in which they were printed in the test blank used in this study.

(1) False premises. In this test, devised by Thurstone,[10] there are 25 nonsense-type syllogisms, each of which consists of two premises and a conclusion to be checked as true or false. The verbal factor explains less than 3% of this test's variance.

(2) Essential operations. This test of arithmetical reasoning requires no number manipulation; only 6% of its variance is explained by the number factor. The task is to indicate, for each of the 10 posed problems, which 1 of 5 given items of information is *not* essential to its solution.

(3) Syllogisms. Each of 5 standard-type syllogisms is followed by 5 conclusions, giving a total of 25 conclusions to be judged true or false. Verbal variance is practically nil in this test, less than 1%.

(4) Problem-solving. As stated by the authors, an attempt was made in devising this test "to reduce to a minimum the number variance by keeping all manipulations simple and by allowing approximate answers."[11] Their success is attested by the resulting number variance, which is less than 1%. The test, however, is not as pure a measure of *logical* reasoning as one would wish—the loading on general reasoning being higher than, and the loading on education of conceptual patterns being almost as high as that on the criterion-factor.[12]

To summarize briefly, the Reasoning Test used as the criterion-measure consists of 4 sub-tests (70 items), selected on the basis of high variance on logical reasoning and low variances on verbal, numerical, and perceptual factors. Administration time is 48 min. and the maximal score possible is 70. Odd-even reliability deter-

---

[8] J. P. Guilford, A. L. Comrey, R. F. Green, and P. R. Christensen, A factor-analytic study of reasoning abilities: I. Hypotheses and description of tests, *ibid.*, 1950, (no. 1), 1-23.

[9] The factor content referred to here and in the discussion following is based on the final rotated matrix-X, Guilford, Green, and Christensen, *op. cit.*, 13.

[10] The writer is indebted to Dr. Thurstone for permission to use this test, and to Dr. Guilford for permission to use the other three tests. The latter supplied copies of, and scoring keys for all the tests.

[11] Guilford, Comrey, Green, and Christensen, *op. cit.*, 9.

[12] Guilford, Green, and Christensen, *op. cit.*, 13.

mined on 488 Stanford undergraduates, the total sample from which the High and Low Groups were drawn, is 0.824—stepped up by the Brown-Spearman formula.

It was expected that the Reasoning Test would tap an ability different from that usually referred to as 'general intelligence.' This was predicated on the conviction that many individuals of high test-intelligence do not excel in problem-solving ability, nor in the ability which we have assumed to be an undeniable aspect of it—logical reasoning. Of interest in this regard are the correlations of reasoning ability, as here measured, with intelligence, as measured by examinations of the Educational Testing Service. For 268 of the first 300 individuals who took the Reasoning Test, scores on both parts of the Aptitude Test were available.[13] The resulting correlations with Reasoning score are 0.39 for Verbal Aptitude and 0.58 for Mathematics Aptitude. Obviously, the Reasoning Test and the Aptitude Test are not measuring the same ability. Mathematics Aptitude is more highly related to Reasoning score, a fact which is entirely consistent with the definition of the criterion variable.

High and Low Groups. The High and Low Groups were selected from 308 men and 180 women students, mostly sophomores, who were registered in classes in general psychology during the academic year 1952-1953 and served as *Ss* as a course requirement. The ratio of 1.7 men to one woman in this sample is typical of the sex ratio for undergraduates at Stanford, which was 1.9 for the year 1952-1953.

The Reasoning Test was administered to groups ranging in size from 40 to 175; score means and sigmas were determined for the sexes separately. Since the difference between the sex means was significant ($CR = 5.2$), the upper 15% and the lower 15% of each sex distribution were selected to make up High and Low Groups. This procedure precluded the possibility that any differences, found between High and Low Groups on the experimental variables, would be a reflection of sex differences.

Means and sigmas for the total sample and for extreme groups are given in Table I.

Free and Controlled Word Association

Research workers in the general area of problem-solving, thinking, and reasoning differ in their convictions as to the type of materials which should be studied, as to which particular function in the process or processes is crucial, and as to what the interrelationships of these functions are. Nevertheless, all seem to agree that at least three functions are important for efficiency in problem-solving. These three, each of which is variously referred to in the literature, are: (1) recognition of, formulation of, or orientation to, the problem; (2) fluency of relevant ideas, variability, or ability to bring past experience to the problem; and (3) a critical, judging, or weighing function. It is with the second function that this section will be mainly concerned.

---

[13] The writer is grateful to E. J. Sweeney, who obtained these scores from the Registrar's Office at Stanford, and computed the correlations reported here.

Johnson, in his excellent review of problem-solving literature to 1943, suggests that mental activity during problem-solving is an interplay of free and controlled association.[14] He stresses the importance of studying individual differences in these processes, and commenting on Murphy's finding that individuals in different academic disciplines differ as regards type of other-word associations,[15] says, "If this is a general difference between these groups in the free play of thought, it will affect their solutions of complex problems."[16]

The fact that group differences in free association can be demonstrated plus the fact that there is a "general and consistent tendency for an individual to give the same type of responses on repeated tests."[17] indicated to the writer that a search for quantitative and qualitative differences between 'good' and 'poor' reasoners, at the very simple level of word-association, should be fruitful. In this paper only the study of quantitative differences is to be reported; a content analysis of responses is in progress and will be published later.

TABLE I

DATA ON REASONING TEST

Group	Men			Women		
	N	*M*	*SD*	*N*	*M*	*SD*
Total sample	308	39.35	13.03	180	32.72	13.88
High	46	58.11	3.68	27	53.89	4.01
Low	46	19.09	6.24	27	11.33	6.55

It was felt that speed of unrestricted word flow would not significantly differentiate the High and Low Groups.[18] Certainly the 'fluid diffuseness' characteristic of some types of psychoses is not highly associated with productivity in the realm of reasoning. It was expected, however, that any requirement restricting the response and necessitating selection from a smaller reserve of relevant responses would on the average be more inhibiting for the Low than for the High Group.

The specific hypothesis to be tested is that good and poor reasoners, selected from a sample which is relatively homogeneous as regards general intelligence, will not differ in speed of completely free word-association but will differ in speed of controlled association, the difference favoring the good reasoners and increasing with each additional restriction placed on the response.

[14] D. M. Johnson, A modern account of problem solving, *Psychol. Bull.*, 41, 1944, 201-229.

[15] Gardner Murphy, An experimental study of literary *vs.* scientific types, this JOURNAL, 28, 1917, 238-263.

[16] Johnson, *op. cit.*, 219.

[17] C. M. Keene, Commonality of response on a word-association test: A study of standardization procedures and an attempt to forecast moderate emotional maladjustment, unpublished doctoral dissertation, Stanford University, 1951, 135.

[18] The terms 'flow' and 'selection' used here are the names given the two factors extracted by D. M. Johnson and Floyd Reynolds, A factor analysis of verbal ability, *Psychol. Rec.*, 4, 1941, 183-195.

Subjects. For this and all other experiments reported in this paper the *Ss* are the High and Low Groups, each of which consists of 73 individuals (46 men and 27 women), selected as described earlier.

Procedure. The original reasoning tests had been numbered and all subsequent tests were so identified by a corresponding case number that *E*, who did all the testing, did not know to which group any *S* belonged until after all tests were scored. The *Ss* knew, of course, that they had been chosen to participate in two further 1-hr. experiments, but none knew on what basis he had been selected.

Ss were tested in groups of four or five, spaced around a large table at which *E* was also seated. All word-association tests were taken at one sitting, with a 2-min. rest-period between tests.

An effort was made to assure *Ss* that no clinical analysis of their responses would be made and to emphasize the requirement of quantity of production even at the expense of handwriting and spelling. They were told to write as fast as they would on an examination which they knew they would not have time to finish but which the instructor must be able to read. Since number of words written in a given time must at least in part be a function of the individual's speed of handwriting, a measure of this speed was obtained by having the *Ss* write the same word, *pearls,* as many times as possible in 3 min. This word was chosen as being of average length and as having in it a variety of letters, among which were the three letters, *p, l,* and *s,* to be repeatedly written in Tests (2) and (3).

The word-association tests proper, and a brief of instructions specific to each, are here listed in the order in which they were given. Three minutes each were allowed for Tests (1) through (3); for Test (4), Synonyms, 13 min. Time required to administer the Kent-Rosanoff test varied in different groups from 12 to 14 min.

Tests:

(1) Any words. Write as many words as possible; just any words will do.
(2) *P—words.* Write as many words as possible which begin with the letter *P.*
(3) *S—L words.* (Three examples were given.) Write as many words as you can which *begin* with *S* and *end* with *L.*
(4) *Similar words.* (Two examples were given. There are eight simple stimulus-words in this test.) Write as many words as you can which have about the same meaning as the word given.[19]
(5) *Kent-Rosanoff Word-Association Test.*[20]

Score for Test (1) was simply the number of words written, for the next three tests it was number of relevant words. Spelling errors were disregarded unless the misspelling served to bring the word within the restriction imposed for the test, *e.g. shale,* misspelled *shal,* was not counted as satisfying the requirements of the S—L test.

---

[19] This test is 'Controlled Association' from Thurstone's Primary Mental Abilities Battery. It is here designated 'Similar Words' to avoid confusion with other tests, Tests 2, 3, and 4, of controlled association. The writer is indebted to the author for permission to use both this test and his First and Last Letters Test, here referred to as 'S-L words.'

[20] The Kent-Rosanoff is listed here since it was administered with Tests 1–4. No time-analysis was made; content analysis of this test and of Test 1 will be reported later.

It will be noted that in Tests (2) and (3) only structural restrictions are involved whereas in Test (4) the restriction is one of meaning. Although the stimulus-words are extremely simple words, like *good* and *happy,* Test (4) is in essence a near-synonyms test and therefore must to some extent tap vocabulary. Offhand the criticism might be made, since vocabulary has been shown to be the best single measure of general intelligence, that any differences obtained on this test would be reflecting differences in intelligence. The criticism loses its potency, however, when one considers certain facts. In Thurstone's Primary Mental Abilities Battery this test correlated only 0.25 with the Thorndike Vocabulary Test. Furthermore, it will be remembered that the test of reasoning ability is composed of tests selected as having negligible saturations on the verbal factor and also that the Reasoning Test correlates only 0.39 with verbal aptitude.

Results and discussion. The results of comparisons of High and Low Groups on the word-association tests and on the speed-of-writing test are presented in Table II.

Attention is called to the unexpected significant difference ($P = 0.01$), in speed of writing, favoring the Low Group. During try-out runs of these tests on a few preliminary *Ss*, large individual differences in speed of writing had been observed.

TABLE II

Comparison of High and Low Groups on Word-Association Tests

Tests	Mean		t	P	s_W	Diff./s_W
	High	Low				
(1) Any words*	67.58	66.65	0.51	.62	11.12	.08
(2) P–words*	36.21	33.40	2.50	.006†	6.63	.42
(3) S–L words	10.97	8.73	3.61	<.0002†	3.75	.60
(4) Similar words	43.60	37.55	5.60	<.00003†	6.52	.93
Speed of writing	82.34	86.10	2.57	.01	8.81	—

* Means for these tests adjusted for speed of writing by analysis of covariance; t=the square root of F.

† P-value for one-tailed test since direction of difference was predicted.

Consequently, a speed-of-writing test was included that allowance could be made for this factor. It is hard to believe that the High Group cannot write as fast as the Low; it seems more likely that the difference found reflects a conative difference—a lesser willingness on the part of the Highs to compromise quality for the sake of quantity.

With regard to the main comparisons in Table II, it should be noted that the means for the first two tests have been adjusted for writing speed by the covariance method whereas means for Tests (3) and (4) have not. Adjustments to the latter were unnecessary since there was no correlation with speed for either test within either group. A glance at the *P*-values in the table will show that while the difference in mean number of completely free word-associations—Test (1)—is not significant, the difference

in speed of controlled word-associations—Tests (2)–(4)—are not only significant but increase in significance as the test-requirements become more restrictive.

These results are entirely in line with the hypothesis but, as the reader will no doubt have realized, the hypothesis is difficult to prove statistically. A relatively insensitive test can be applied—the probability of four tests yielding differences which increases in the order predicted is 0.04. Perhaps more meaningful are the ratios given in the last column of Table II, where the difference between means for each test is expressed relative to the respective within-groups variation. The resulting ratios, which like standard scores are directly comparable, increase with each successive test.

It might be suspected that this increase in differentiability is merely reflecting an increase in reliability from test to test. The following coefficients, determined on a comparable group of 199 *Ss*, proves this not to be the case: Test (1), 0.76 (test-retest, after an 8-wk. interval); Test (2), 0.70 (B — vs. D—words); Test (3), 0.58 (S—L vs. T—E words); Test (4), 0.87 (odd-even item).

The fact that differentiation between good and poor reasoners increases with the degree of restriction placed on the response strongly suggests that the important variable, at least at the simple word level, is speed of *selection* of appropriate responses. It would be worthwhile to learn whether similar results could be obtained with less simple media than isolated words, say with ideas, plans of action, concepts, and hypotheses. It was inferred by Johnson and Reynolds that "flow of relations, concepts, schemas, hypotheses, insights, plans of action, et cetera" might be included in the "flow of response" factor which they extracted;[21] however, Taylor was led by the results of his factor analysis of fluency to believe that "the flow of response in word fluency tasks and the flow of response in ideational fluency are not the same basic process."[22]

DEDUCTION AND INDUCTION

The results of an attempt to study experimentally the use and efficacy of nine 'known methods of attack' on reasoning problems are reported in a recent monograph by Burack.[23] The problems used included a deduction- and an induction-test, both borrowed from Thurstone's early reasoning bat-

[21] Johnson and Reynolds, *op. cit.*, 192.

[22] C. W. Taylor, A factorial study of fluency in writing, *Psychometrika*, 12, 1947, 239-262.

[23] Benjamin Burack, The nature and efficacy of methods of attack on reasoning problems, *Psychol. Monog.*, 64, 1950, (no. 313), 1-26.

teries, and a geometrical figure problem. Two of Burack's conclusions, especially pertinent here, were in effect: (1) there are wide individual differences in the extent of use of a particular method; and (2) the extent of use and the efficacy of a particular method is dependent upon the kind of task involved.

Although the evidence for the effectiveness of the various methods was not conclusive, it seemed to the writer that methods of arriving at solutions to problems of the deductive and inductive type should differentiate good from poor reasoners. Consequently, it was decided to include in the present experiment the same deduction- and induction-tests used by Burack,[24] and to follow up with questions suggested by this author to test certain methods. Six methods were selected from Burack's nine as most promising for our purpose: (1) preliminary survey of all aspects of the material; (2) application of past experience; (3) varied trials; (4) control; (5) locating a crucial aspect of the problem; and (6) elimination of sources of error. An additional aim of this part of the experiment was to determine whether or not the High and Low Groups would differ in time to solve problems.

Since deduction- and induction-tests are reasoning tests, it goes without saying that accuracy scores on both tests should differentiate groups which have been selected as good and poor reasoners. It was also expected that all differences in responses to questions of method would favor the High Group. No prediction was made, however, as to which group would be faster in solving the problems.

Procedure. Three tests were given in a single session: the deduction- and induction-tests, which are of immediate concern here, and a volume-measuring series of the Luchins' type, to be considered in the next section of this paper.

Although unlimited time was to be allowed, a record was to be kept of the time required by every *S* to complete each test. To facilitate this, the *Ss* were tested in groups of 4 or less. Of the 146 *Ss*, 132 were tested in groups of 4; 9, in groups of 3; and 2, singly. General instructions made clear that, in contrast to the earlier word-association session when the emphasis had been on speed, in this session the important factor would be accuracy.

Deduction-test. The deduction-test is reproduced here in the form used in the experiment.

Deduction-problem: (a) Captain Watts and his son James have been found shot—the father in the chest and the son in the back. (b) Both clearly died instantaneously. (c) A gun fired close to the person—as, for example, when a man shoots himself—will blacken and even burn the skin or clothes; fired from a greater distance it will leave no such mark. (d) The two bodies were found near the middle of a large hall used as a rifle range. (e) Its floor is covered with damp sand which shows every footprint distinctly. (f) Inside the room there are two pairs of footprints only. (g) A third man standing just outside the door or window could aim at any part of the room, but the pavement outside would show no footmarks. (h) Under Captain Watts' body was found a gun; no such weapon was found near James. (i) In each case the coat, where the bullet entered, was blackened with gunpowder, and the cloth a little singed. (j) Captain Watts was devoted to his son and would have died sooner than harm him purposely; hence it is impossible

[24] The writer is further indebted to Dr. Thurstone for permission to use these tests.

to suppose that he killed him deliberately, even in self-defense. (k) But some think that James secretly disliked his father and hoped to inherit his fortune at his death.

Questions:

(*a*) To what was Captain Watts' death due? Murder ——— Accident ——— Suicide ———

(*b*) To what was James' death due? Murder ——— Accident ——— Suicide ———

Specific instructions for this test were to read the paragraph carefully as many times as desired, answer Questions (*a*) and (*b*), and hand the paper immediately to *E*. When all in the group had finished, *S* was given a sheet on which were printed the six questions used to probe for methods. To minimize misunderstandings, *E* read each question aloud with additional explanations and *S* wrote his answer in the space provided. A copy of the questions asked will be found under the section on results, where their inclusion will expedite the discussion.

Induction-test. The induction-test is a letter-grouping test involving concept formation. In the form used in this study there are 2 examples, 4 practice problems, and 19 test items. The task is to discover, for each item of five 5-letter groups, that group which does *not* contain the feature common to the other four groups. One of the practice problems is given here for the purpose of illustration; the reader is referred to Burack's monograph for a reproduction of the entire test.[25]

ABCDE XYZPS EFGHI CXVTN PJKLM

Four of the groups have letters occurring in alphabetical order; therefore, the fourth group, which does not, should be underlined.

A word should be said about the order of administration of the three tests in this session. The deduction-problem was given first and the induction-test last since it was expected that variability in time required would be least for the former and greatest for the latter test. Furthermore, as the induction-test required no specific oral instructions, *S* could easily be put to work on the last test (as soon as he had finished the volume-measuring series) without disturbing the others in his group. Because of time limitations only one of Burack's questions regarding method could be included for the induction-test.

Results and discussion. The comparisons of High and Low Groups with respect to both accuracy and speed on the deduction- and induction-tests, will be found in Table III. It should be noted that whereas mean accuracy scores are compared for the induction-test, the analogous comparison for the deduction-problem involves the percentages of the two groups answering Questions (*a*) and (*b*) correctly, *i.e.* checking the correct cause for both deaths.

As predicted, the better reasoners are more accurate even when there is no premium on speed; the difference for each problem is significant beyond

[25] Burack, *op. cit.*, 3.

the 0.00003-level. Comparisons of time-scores yield differences in means and standard deviations which are, however, in opposite directions for the two tests. As can be seen from the table, the High Group is, on the average, slightly slower ($P = 0.10$) in solving the deduction-problem but they are significantly faster ($P < 0.002$) in completing the letter-grouping series. One is tempted to suggest that the nature of the particular deduction-problem used here is such that *S* could be superficial reading of the paragraph and selection of a few details make a plausible decision as to the

TABLE III

COMPARISON OF HIGH AND LOW GROUPS WITH REGARD TO ACCURACY AND TIME-SCORES ON THE DEDUCTION- AND INDUCTION-TESTS

	Test	N		Mean		SD		CR	P
		H	L	H	L	H	L		
Accuracy	Deduction	73	73	79%	41%	—	—	4.74	<.00003*
	Induction	73	73	17.55	15.73	1.53	2.00	6.18	<.00003*
Time (min.)	Deduction	73	73	3.59	3.22	1.50	1.16	1.67	.100
	Induction	73	73	19.42	23.15	6.19	7.73	3.22	<.002

* P-value for one-tailed test since direction of difference was predicted.

causes of the two deaths. The nature of the induction-problem, on the other hand, is such that *S* must actually examine every letter before he can discover the feature common to four groups but lacking in the fifth. This speculation is consistent with our earlier interpretation of the difference in handwriting speed as a reflection of conative difference between good and poor reasoners.

To facilitate the discussion of results of the methods analysis, the questions are listed here exactly as they were printed on a sheet given to *S* *after* he had made his decision as to the causes of the two deaths. The only question asked after completion of the induction-problem is also listed here.

Questions on deduction-problem:

(1) How many times did you read the entire paragraph? ———

(2) (a) Have you ever had any previous experience with a similar type of problem? Yes ——— No ———
(b) If answer is 'yes,' were you able to apply this past experience here? Yes ——— No ———

(3) Did you try to apply the possibility of *each* of the three causes—*murder, accident, suicide*—to the death of Captain Watts, and then to the death of James, to determine the correct cause for each death? Yes ——— No ———

(4) Did you single out *each sentence* to try to determine whether information in any of the various sentences made *impossible* any of the three causes? Yes ——— No ———

(5) Did you conclude that certain bits of given information were important or relevant for solving this problem? If so, list all such points. You may consult the paragraph, but list only those points which you actually used in determining the cause of each death. List letters only.

(6) Answer the following Questions, (*a*)—(*d*). When answer is 'yes,' write down the letter (or letters) indicating the information.

(*a*) Is there any information in the paragraph which makes the following sentence *not* helpful in solving the deaths: "A third man standing just outside the door or window could aim at any part of the room but the pavement outside would show no footprints."

(*b*) Is there any information to indicate who died first—Captain Watts or James?

(*c*) Is there any information which makes murder improbable for James' death?

(*d*) Is there any information which makes suicide improbable for James' death?

Question on induction-problem:

(1) List all the ways you can remember in which four groups had something in common that was lacking in the fifth group.

A summary of the results is given in Table IV. Note that the proportions of High and Low Groups making a given response are compared for the first four questions on the deduction-test, and for the only question on the induction-test. For Questions (5) and (6), however, mean values are compared. The score for an individual on Question (6) was the total number of correct items listed for sub-parts (*a*) through (*d*). In all instances *P*-values are for one-tailed tests since direction was predicted for all differences.

TABLE IV

COMPARISON OF HIGH AND LOW GROUPS WITH REGARD TO METHODS USED IN DEDUCTION- AND INDUCTION-TESTS

	Question no.	Response	Method tested	Base N H	Base N L	Proportion H	Proportion L	CR	P
Deduction-test	(1)	Read more than once	Preliminary survey	73	73	.52	.47	.66	.26
	(2a)	Yes	—	73	73	.49	.25	3.99	.001
	(2b)	Yes	Application of past experience	36	18	.33	.40	.40	.66*
	(3)	Yes	Varied trials	73	73	.67	.68	.18	.57*
	(4)	Yes	Control	73	73	.88	.75	1.92	.03
	(5)	No. points listed	Locating a crucial aspect	73	73	6.79†	5.14†	6.02	<.00003
	(6)	No. items listed	Elimination of sources of error	73	73	4.21†	2.60†	5.90	<.00003
Ind. T.	(1)	No. ways listed	Locating a crucial aspect	73	73	.70	.41	3.50	.0002

* *P*-value = probability of a greater difference in the predicted direction.
† Mean values.

Where a difference is in the direction opposite from that predicted, the *P*-value is to be interpreted as the probability of a greater difference in the expected direction.

These results are not impressive. Contrary to expectation, differences are totally insignificant for Questions (1), (2b), and (3) designed to test the three methods: preliminary survey, application of past experience, and varied trials. That these methods should differentiate good from poor reasoners still seem a reasonable expectation in spite of the results reported

here but, after careful reconsideration of the questions used, it is our opinion that the latter are inadequate for testing these methods. Conceivably one critical reading, or one reading plus rereading of only crucial parts, might constitute a better preliminary survey than several less critical readings. This conjecture is given some support by the already cited fact that good reasoners are significantly more accurate but slightly slower than poor reasoners on this problem. With regard to varied trials, while on the average and for most problems it may be true that the more varied the trials the more probable is solution, it might also be that in some situations trying every possibility would be mere compulsiveness. Question (*3*) forces a 'no' answer unless *each* cause has been applied to *each* death.

Although application of past experience as here determined fails to differentiate the two groups, significantly more of the High Group ($P = 0.001$) said they had had previous experience with this type of problem. Any or all of three factors may be operating to produce this difference: (1) real difference in experience; (2) better recognition, on the part of the High Group, of similarity between this problem and others in their experience; or (3) failure to admit such experience—rationalization for poor performance—on the part of the Low Group. Maier has argued against the use of introspective techniques in studying processes underlying problem-solving on the ground that reports of vague experiences are not only incomplete but are apt to be loaded with rationalizations.[26] Application of past experience, insofar as one may be unaware of such application, will be incompletely reported.

Furthermore, the control method might have yielded a more significant difference than that found ($P = 0.03$) had the question been a more adequate test of control. It seems in retrospect that, since two of the sentences, (g) and (k), in the problem are irrelevant to the solution, some of the better reasoners might not have "singled out *each* sentence" as required for an affirmative answer to Question (*4*).

As indicated by the responses to Questions (*5*) and (*6*) the High Group far excels the Low both in locating a crucial aspect and in eliminating sources of error. The former method also discriminates between the High and Low Groups ($P = 0.0002$) on the induction-test.[27] These results

---

[26] N. R. F. Maier, The behavior mechanisms concerned with problem solving, *Psychol. Rev.*, 47, 1940, 43-58.

[27] The question asked here to test "locating a crucial aspect" was used by Burack, (*op. cit.*, 10) to test "analysis into major variables," which he admits overlaps with the former method (*ibid.*, 2). We regard each 'way' listed as a crucial aspect located.

are not inconsistent with those reported by Burack. Though his statistical treatment makes evaluation of his results difficult, it is safe to say that the two methods, locating a crucial aspect and elimination of sources of error, were more highly related than other methods to success in problem-solving. Unquestionably, good reasoners are better at selecting relevant and eliminating irrelevant aspects of a problem. This is further evidence of the importance of *selection* as a variable in reasoning and problem-solving.

The negative character of some of the results in this experiment has in no way lessened the writer's conviction that methods of problem-solving should differentiate good from poor reasoners, but it has necessitated a reappraisal of the technique used. More careful attention to the selection of probing questions is definitely indicated. Perhaps if individuals could be trained to verbalize every thought as they solved problems and if a recording device were used, one might better determine the efficiency of problem-solving methods. That this procedure would be feasible is suggested by Patrick's success in her study of the creative process in poets and artists.[28]

Experimentally Induced Set

When the present study was first outlined, it was planned that the two groups of reasoners should be given one of the Luchins' volume-measuring series as a test of behavioral rigidity. This intention was reinforced by Guetzkow's stimulating analysis of the operation of set in problem-solving behavior.[29] Of particular merit is his delineation of the two processes, "susceptivity to set" and "ability to surmount set," a distinction which very recently has also been made explicit by Luchins.[30] In our opinion this distinction was implicit in Luchins' earlier articles.[31] It is to be hoped that the emphasis placed by both these authors on the disparateness of the two processes will serve to bring order out of the confusion prevalent in current psychological literature concerned with 'set' or 'rigidity' phenomena.

Oliver and Ferguson have distinguished two types of 'rigidity' tests,

---

[28] Catherine Patrick, Creative thought in poets, *Archiv. Psychol.*, 26, 1935, (no. 178), 1-74.

[29] Harold Guetzkow, An analysis of the operation of set in problem-solving behavior, *J. Gen. Psychol.*, 45, 1951, 219-244.

[30] A. S. Luchins, The Einstellung test of rigidity: Its relation to concreteness of thinking, *J. Consult. Psychol.*, 15, 1951, 303-310.

[31] Luchins, Proposed methods of studying degrees of rigidity in behavior, *J. Personality*, 15, 1947, 242-246; Rigidity and ethnocentrism: A critique, *ibid.*, 17, 1949, 449-466.

which they purport to show by factor analysis are unrelated:[32] (1) tests "demanding reorganization of culturally induced behavior patterns," and (2) tests "demanding reorganization of experimentally induced behavior patterns." The variable to be investigated in this study is designated *experimentally induced set,* operationally defined as the tendency to persist in an experimentally established mode of response in situations where that mode of response is no longer the most efficient, or most direct, method.

The main hypothesis to be tested is twofold. Good and poor reasoners will not differ in susceptibility to experimentally induced set, but will differ in ability to overcome set thus established. A further hypothesis, which can be tested with data on the same problem, and which will become clearer to the reader when he has read the sections on Procedure and Results, concerns Guetzkow's conclusion that ability to solve those problems, used in the volume-measuring series to induce set, is not related to reasoning ability. We were not interested in the question as to whether such problems require productive or merely reproductive reasoning, but we were convinced that solvers and non-solvers of even the simplest arithmetic problems should be differentiated by scores on a reasoning test. Therefore, it was predicted that good and poor reasoners would differ significantly in ability to solve the five 'set' problems.

Procedure. The volume-measuring series used in the present analysis is exactly the same as that used by Guetzkow. This was patterned after the typical Luchins series of five 'Einstellung,' four 'critical,' and four 'extinction' problems but with the number of critical and extinction-problems reduced to two and one, respectively. One fundamental difference in our procedure should be noted. Whereas Luchins and Guetzkow presented and timed the problems individually, here two examples and the eight problems were so printed on a single test sheet that *S* could work through the list of problems at his own speed though he was never permitted to go back to an earlier problem. This test sheet is reproduced here to facilitate exposition.[33]

Water-Jar Test:

In each of the following problems you are to suppose that you have available three water-jar containers of different volumes and you are to obtain a volume of water which is different from the volumes of the three original containers. The water supply is unlimited.

Examples:

(1) Jars available: 4 quart, 10 quart, and 3 quart. You are to measure out

---

[32] J. A. Oliver and G. A. Ferguson, A factorial study of tests of rigidity, *Canadian J. Psychol.,* 5, 1951, 49-59.

[33] The title and the examples for this test were borrowed from the mimeographed material distributed at the APA, Sept., 1952, by M. Wiener, E. L. Cowen, and J. H. Hess, Generalization of problem solving rigidity.

exactly 7 quarts. *Answer:* Fill the 10 quart, pour off 3 quarts. 10 − 3 = 7; or fill the 4 quart and the 3 quart. 4 + 3 = 7.

(2) Jars available: 42 quart, 71 quart, and 5 quart. Measure out 19 quarts. Answer: Fill the 71 quart, pour off 42 quarts, pour off 5 quarts, pour off 5 quarts again. 71 − 42 − 5 − 5 = 19.

Problems: Indicate your answer for each as indicated in the examples.

Number	Jars available	To obtain	Solution
(1)	21, 127, 3	100	——
(2)	14, 163, 25	99	——
(3)	18, 43, 10	5	——
(4)	9, 42, 6	21	——
(5)	20, 59, 4	31	——
(6)	23, 49, 3	20	——
(7)	15, 39, 3	18	——
(8)	28, 76, 3	25	——

In the discussion to follow, the three jars will be referred to as A, B, and C in the order listed for each problem. In all Luchins' series, the set problems are solvable only by the formula B − A − 2C; the criticals, which test the strength of set, may be solved either by this same formula or by a more direct method, *e.g.* A + C or A − C; and the extinction-problems, which indicate ability to overcome set, are solvable only by one of the direct methods. It was discovered, but unfortunately not until the results were being tabulated, that the fourth 'set' problem in this series is solvable not only by the B − A − 2C formula but also by a more direct formula, A + 2C. It is believed that this fact may not have been noted by Guetzkow and that failure to take it into account may explain his finding no difference,* as regards reasoning score, between those solving and those failing to solve the five 'set' problems.

Results and discussion. The results of all comparisons of High and Low Groups on the Water-Jar test may be found in Table V, where in the second and third columns the letter *F* refers to the Formula B − A − 2C and *D* refers to the direct method: A + C or A − C. It will be seen that 68%, or 50 *Ss* in each group, solved Problems (1) to (5) by the B − A − 2C formula. This percentage is very close to the 70% reported by Guetzkow for his total group. Thus analyzed, our data would corroborate Guetzkow's conclusion that ability to solve the first five problems is unrelated to reasoning ability. When, however, the 15 *Ss* in the High Group and the 7 *Ss* in the Low Group who failed to solve only Problem (4) by the B − A − 2C method—simply because they solved it correctly by the more direct A + 2C method—are treated as solving all five problems, the percentages of the High and Low Groups become 89 and 78, respectively. This difference is significant at the 0.05-level. If Guetzkow did not note such *Ss* in his group —there must have been some who solved Problem (4) directly—he un-

* We have written Dr. Guetzkow, calling his attention to the nature of Problem (4), but to the present time have had no reply; therefore, deductions made herein concerning Guetzkow's treatment are merely assumptions on our part.

doubtedly included some of his better reasoners in his 'failed to solve' group. Further evidence of the relationship of problem-solving to reasoning ability is afforded by comparing the percentage in each group who solve correctly all eight problems in the series, whether by the direct method or by the longer formula. This difference, given in the last line of Table V, is significant beyond the 0.00003-level. It seems safe to say that the hypothesis that ability to solve problems in the volume-measuring series would differentiate good from poor reasoners is confirmed.

Following Luchins' procedure, we have included only those Ss who solved all five problems by the B − A − 2C formula, *i.e.* those who have

TABLE V

COMPARISON OF HIGH AND LOW GROUPS WITH REGARD TO SOLUTIONS TO THE WATER-JAR PROBLEM

Problems solved	Method used*	Ss included in comparison	Base N		Percentage solving		CR	P
			H	L	H	L		
(1)–(5)	*F*	All	73	73	68	68	.00	.50†
(1)–(5)	*F*+(4)*D*	All	73	73	89	78	1.89	.03†
(6) and (7)	*D*	(1–5) *F*	50	50	26	16	1.33	.19
(8)	*D*	(1–7) *F*	58	54	84	38	3.49	.0001†
(8)	*D*	(1–6) *F*	50	50	90	53	4.10	<.00003†
(1)–(8)	D or *F*	All	73	73	79	41	4.97	<.00003†

* In this and the next column, *F* refers to the method of Formula B − A − 2C and *D* refers to the Direct method, *i.e.* A + C or A − C.

† P-value for one-tailed test since direction of difference was predicted.

had equal practice in solving problems by the set-inducing method, in the analysis of subsequent problems. A glance at the comparison in the third line of the table will show that the difference between High and Low Groups, as regards susceptibility to set, is not significant. It should be noted that 26 and 16 are the percentages solving both criticals, Problems (6) and (7), by the direct formula.

To test the hypothesis that the reasoning groups would differ in ability to overcome set, two comparisons were made. In the first only those Ss solving all seven problems by the B − A − 2C method are included. If, however, one is interested in those evidencing ability to overcome set, it should be legitimate to include all those who solve Problem (8) directly, regardless of the method used for Problems (6) and (7). Although we are in agreement with Luchins' objection to the use of the critical problems alone as a criterion of overcoming set,[33] it seems unfair to exclude individuals who satisfy the accepted criterion, solving Problem (8) directly, simply

[33] Luchins, 1951, *op. cit.*, 308.

because they had overcome set on the preceding problems. These two comparisons yield differences which are significant at the 0.002-level and beyond the 0.00003-level, respectively. Obviously, the ability to overcome experimentally induced set differentiates good from poor reasoners.

The failure of *susceptibility to set* to yield a significant difference between groups which are significantly differentiated by the *ability to overcome set* is entirely consistent with results reported by Guetzkow,[36] by Luchins,[37] and by Adamson and Taylor.[38]

Summary and Conclusions

This study was designed to determine the relationship of selected variables to logical reasoning ability within a college sample. A criterional test was used to select groups high and low in this ability.

Conclusions warranted by the data, which were obtained to test specific hypotheses concerning the discriminability of the variables studied, may be summarized as follows.

Speed of word-association. High and Low Groups do not differ in speed of unrestricted word flow, but do differ in speed of controlled association; the difference, favoring the High Group, increases in significance with each restriction placed on the response.

Deduction and induction. (1) Within the limitations of the particular tests used to measure these processes, it is concluded that good reasoners are superior as regards accuracy in both processes and as regards speed of induction. In speed of deduction the groups do not differ. (2) Of six methods of reasoning studied—preliminary survey, application of past experience, varied trials, control, locating a crucial aspect, and elimination of sources of error—only the last three yield significant differences between High and Low Groups. Inadequacies of the technique used to get at methods are discussed.

Experimentally induced set. (1) Contrary to the results reported by Guetzkow, the ability to solve problems involved in the volume-measuring series is here found to differentiate good and poor reasoners. A possible explanation for the discrepancy in results of the two studies is suggested. (2) Good and poor reasoners differ significantly in the *ability to overcome set* but differ little in *susceptibility to set*—a finding which is consistent with the results of Guetzkow and other investigators.

---

[36] Guetzkow, *op. cit.*, 223.
[37] Luchins, 1951, *op. cit.*
[38] R. E. Adamson and D. W. Taylor, Functional fixedness as related to elapsed time and set, *J. Exper. Psychol.*, 47, 1954, 122-126.

A COMPARISON OF UNIOCULAR AND BINOCULAR CRITICAL FLICKER FREQUENCIES: SIMULTANEOUS AND ALTERNATE FLASHES

By GARTH J. THOMAS, University of Chicago

The nature of the coupling mechanism combining the two uniocular inputs into a fused, unitary cyclopean image remains largely a matter of speculation. On the basis of what he interpreted as negative results from his experiments on binocular flicker, Sherrington concluded that binocular fusion was *not* based upon a neural mechanism analogous to that underlying the integration of impulses funneling into the motor, final, common path.[1] He compared binocular critical fusion frequencies (*CFF*) when the train of flashes was delivered synchronously to corresponding (*i.e.* fused) retinal areas in the two eyes with *CFF* when the flashes were delivered alternately or out of phase. Simultaneous flashes in the eyes yielded what he considered only an insignificant increase in sensitivity to flicker (binocular summation) over that obtaining when the two eyes were stimulated alternately.

Relatively few investigations of this problem of binocular interaction in respect of flicker-discrimination have appeared since Sherrington's classical study.[2] In recent years four studies have appeared, all of which report significant effects of phase relations of the flashes delivered to the two eyes in binocular *CFF*.[3] Crozier and Wolf, in a study comparing uniocular and binocular *CFF*, have emphasized the importance for theoretical conclusions of making such comparisons for the whole *F*—log *I* contour rather than

* Accepted for publication January 14, 1954. This study was supported by the Medical Research and Development Board, Office of the Surgeon General, Department of the Army, Contract No. DA-49-007-MD-36. The author is indebted to Mr. Theodore Schaeffer for assistance in it.

[1] C. S. Sherrington, On binocular flicker and the correlation of activity of 'corresponding' retinal points, *Brit. J. Psychol.*, 1, 1904, 26-59; and *The Integrative Action of the Nervous System*, 1906, 354-386.

[2] For a summary of experiments see G. J. Thomas, Effects of interocular differences in intensity and phase of flash on binocular *CFF*, this JOURNAL, 67, 1954, 632-633.

[3] B. O'Brien and F. H. Perrin, An experimental study of binocular flicker, *J. Opt. Soc. Amer.*, 27, 1937, 63 (Abstract only); F. H. Ireland, A comparison of critical flicker frequencies under conditions of monocular and binocular stimulation, *J. Exper. Psychol.*, 40, 1950, 282-286; C. H. Baker, The dependence of binocular fusion on timing of peripheral stimuli and on central processes: 1. Symmetrical flicker; 2. Asymmetrical flicker; 3. Critical flicker, *Can. J. Psychol.*, 6, 1952, 1-10, 84-91, 123-130, 151-163; and Thomas, *op. cit.*, 632-633.

the more usual procedure of making comparisons at one or at best a few levels of flash luminance.[4] The experiments described in this report compare the effects on binocular flicker contours of synchronous (in phase) and alternate (out of phase) presentation of the flashes to the two eyes in relation to the component uniocular flicker contours.

Apparatus and Procedure

Apparatus. The apparatus has been previously described in detail.[5] Its essential feature may be briefly indicated (see Fig. 1). *S* views two stimuli, one with each eye, through low power microscopes. The stimuli are circular apertures in front of

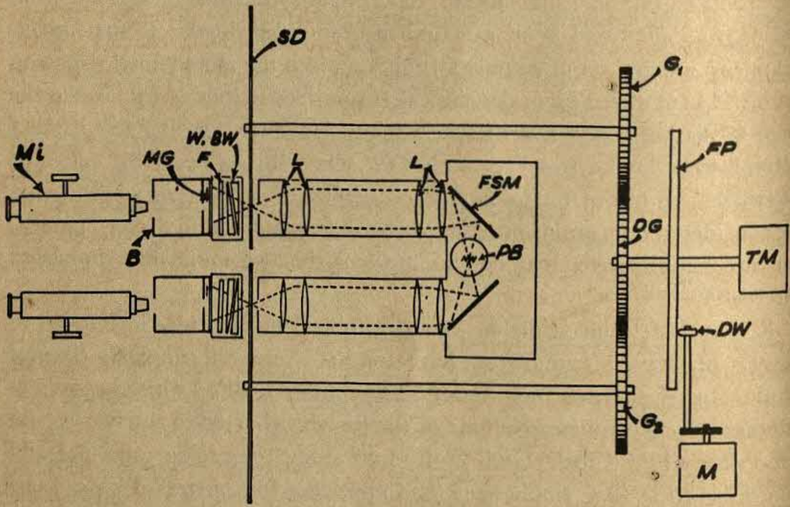


Fig. 1. Schematic Diagram of the Apparatus

Mi, microscopes; *B* box containing thin sheet of glass 45° to the line of sight to reflect fixation-point into the field of view; *MG*, milk glass screens; *F*, *W*, and *BW*, filters, optical wedge and balancing wedge; *SD*, sector disk; *Ls*, lenses; *Gs*, gears; *DG*, driving gear; *FP*, friction plate; *DW*, leather driving wheel; *M*, synchronous motor; *TM*, tachometer magneto.

transilluminated milk glass screens. The luminance of the stimuli can be independently controlled by means of neutral tint optical wedges, balancers, and filters. The microscopes are parallel and adjustable as to separation in order to accommodate individual differences in interpupillary distance. The light beams are interrupted at a point of focus by sector disks rotating on separate but synchronously revolving axles. A fixation-stimulus is projected into the center of each stimulus-patch by reflection of a point of red light from microscopic cover slips mounted 45° to the

[4] W. J. Crozier and Ernst Wolf, Theory and measurement of visual mechanisms: IV. Critical intensities for visual flicker, monocular and binocular, *J. Gen. Physiol.*, 24, 1941, 505-535.

[5] Thomas, *op. cit.*, 634-635.

line of sight between the object lens of the microscope and the stimulus-aperture. When *S* looks through the microscopes, a 6° circular image with a small red fixation-point in its center is projected into each eye, and when the eyes are parallel and accommodated for infinity, the stimuli fall on corresponding retinal areas and appear as a single circular bright area in an otherwise dark field. Care was taken so to align *S*'s head that the exit pupils of the microscopes were centered with respect to the pupils of *S*'s eyes. The head was then held rigidly in place by a biting board—a wax impression of *S*'s teeth.

Procedure. The equipment did not provide a common control of the luminance of both stimuli. To present flashes of equal luminance in the two eyes it was necessary to match the two stimuli at each luminance-level throughout the range to be covered. This was accomplished in the following manner: first, the flash rate was set at a fixed value and right and left uniocular flicker thresholds were separately determined in terms of critical flash luminance. Beginning about 0.2 log unit (millilambert) below threshold, flash luminance was increased by small steps at *S*'s signal until the critical flash luminance was reached and *S* signaled his detection of flicker. *S* signaled by pressing microswitches held in his hands which activated signal lights on *E*'s control panel. This procedure allows *S* to control the stimulus by means of signals without movements gross enough to disturb the alignment of his head and eyes with regard to the exit pupils of the apparatus. He observed the stimulus (maintaining fixation) until he was satisfied it was not flickering, then signaled for a further increase in flash luminance which *E* accomplished by moving the wedge a small amount; and so on until threshold was reached.

After 10 measures of critical flash luminance for each eye had been accomplished, average of the settings in millimeters along the wedge was determined. The wedges for right and left beams were then set at the respective means, and critical flash rates for flicker were determined with each eye alone and under two conditions of binocular regard, flashes in phase and flashes out of phase. The procedure for determining thresholds was the same as that described above except now, the independent variable was flash frequency rather than flash luminance. The procedure was then repeated at other levels until flicker contours had been determined over a range of about six log units. The four uniocular and two binocular thresholds (each threshold involved 10 settings) were determined for a given level in a single session. The uniocular thresholds in terms of critical flash luminance were always measured first in any session, but the uniocular and binocular *CFF* were taken in different orders in different sessions.

Subjects. Two *S*s, Mr. Theodore Schaeffer (*TS*) and the author (*GT*) served in all of the experiments.

RESULTS

Fig. 2 (see also Table I) presents the flicker contours determined in terms of critical flash luminance. The curve for *GT* has been displaced one log unit to the right on the abscissa. Each point is the mean of 10 determinations of critical flash luminance for flicker. Open circles represent thresholds of the left eye; solid circles, the right eye. The curves drawn through the data were fitted by inspection.

TABLE I

MEAN CRITICAL FLASH LUMINANCE (*CFL*) AND VARIABILITY (*SD*) AS A FUNCTION OF FLASH FREQUENCY (*F*) FOR EACH EYE OF EACH S
(*CFL* and *SD* expressed in Log Millilamberts)

Results of *TS*

F	Right eye		Left eye	
	CFL	*SD*	*CFL*	*SD*
9	−2.47	−3.40	−2.79	−3.40
12	−1.88	−2.75	−1.68	−2.39
15	−0.81	−1.78	−0.76	−1.18
18	−0.41	−1.59	−0.51	−1.33
21	−0.25	−1.29	−0.21	−1.14
24	0.01	−1.28	−0.05	−1.34
27	0.16	−1.33	0.18	−1.15
30	0.68	−0.37	0.54	−0.93
33	0.84	−0.40	0.80	−0.52
36	1.07	−0.42	0.97	−0.18
39	1.39	0.07	1.27	0.00
42	1.60	0.58	1.55	0.16
45	1.75	0.45	1.80	0.62
48	2.10	0.72	2.09	0.66
51	2.29	1.02	2.26	1.32
54	2.63	1.60	2.75	1.72

Results of *GT*

F	Right eye		Left eye	
	CFL	*SD*	*CFL*	*SD*
8	−2.46	−3.07	−2.46	−3.03
10	−2.20	−2.83	−2.38	−3.20
12	−1.35	−2.32	−1.44	−2.57
14	−0.71	−1.52	−0.90	−2.06
15	−0.69	−1.66	−0.73	−2.09
16	−0.40	−1.67	−0.42	−1.36
18	−0.21	−1.37	−0.18	−1.61
21	−0.08	−1.15	−0.09	−1.15
24	0.13	−1.11	0.19	−0.97
27	0.41	−0.67	0.41	−0.80
30	0.75	−0.42	0.65	−0.21
33	0.89	−0.28	0.94	−0.14
36	1.37	0.01	1.25	−0.03
39	1.43	0.17	1.44	0.31
42	1.68	0.63	1.64	0.51
45	2.00	0.79	1.96	0.40
48	2.21	1.19	2.23	1.29
51	2.49	1.34	2.67	1.61
54	2.91	1.91	2.88	1.90
57	3.14	1.89	3.15	2.29

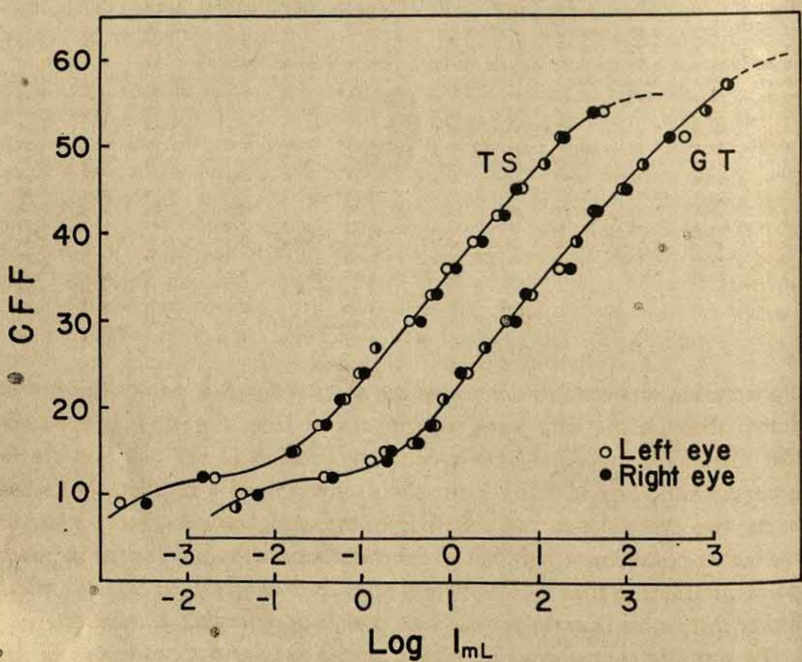


FIG. 2. FLICKER CONTOURS IN TERMS OF CRITICAL FLASH LUMINANCE IN LOG MILLILAMBERTS AS A FUNCTION OF FLASH FREQUENCY
(Contour for *GT* has been displaced one log unit to the right.)

Fig. 3 (see also Table II) shows the flicker contours measured in terms of critical flash rate. The curves drawn through the data were traced from Fig. 2. They make an equally good fit for the data in Fig. 3, indicating that the two methods of measuring flicker thresholds yield very similar contours.

Results for the two conditions with binocular regard, flashes in phase and out of phase in the two eyes, are shown in Fig. 4 (see also Table III). It can be seen from Fig. 2 that at the same flash rate the two eyes did not always yield identical critical flash intensities. In plotting the data shown in Fig. 4, if there was any difference between the two eyes in mean critical flash luminance, as estimated from calibration curves of the wedges, the

TABLE II

MEAN CRITICAL FLICKER FREQUENCY (*CFF*) AND VARIABILITY (*SD*) IN FLASH FREQUENCY PER SECOND (F/SEC.) AS A FUNCTION OF FLASH LUMINANCE IN LOG MILLILAMBERTS (LOG *L*) FOR EACH *S* AND EACH EYE SEPARATELY

Results of *TS*						Results of *GT*					
Right eye			Left eye			Right eye			Left eye		
Log *L*	*CFF*	*SD*	Log *L*	*CFF*	*SD*	Log *L*	*CFF*	*SD*	Log *L*	*CFF*	*SD*
−2.47	8.7	0.50	−2.79	8.4	0.45	−2.46	8.0	0.55	−2.46	8.2	0.23
−1.88	11.3	0.52	−1.68	11.6	0.50	−2.20	9.9	0.30	−2.38	9.9	0.32
−0.81	14.9	0.65	−0.76	14.9	0.62	−1.35	11.9	0.44	−1.44	12.1	0.27
−0.41	18.2	0.75	−0.51	18.3	0.50	−0.71	14.0	0.47	−0.90	14.1	0.27
−0.25	21.2	0.56	−0.21	21.2	0.48	−0.69	14.9	0.30	−0.73	15.1	0.23
0.01	23.9	0.62	−0.05	23.9	0.40	−0.40	16.3	0.33	−0.42	15.9	0.30
0.16	26.6	0.50	0.18	26.4	0.56	−0.21	19.1	0.45	−0.18	18.2	0.45
0.68	30.5	0.50	0.54	30.2	0.48	−0.08	20.7	0.50	−0.09	21.5	0.34
0.84	33.0	0.47	0.80	33.0	0.61	0.13	23.6	0.38	0.19	24.5	0.50
1.07	35.9	0.40	0.97	35.0	0.37	0.41	27.2	0.48	0.41	27.8	0.40
1.39	39.5	0.50	1.27	39.5	0.37	0.75	31.3	0.68	0.65	30.2	0.57
1.60	42.1	0.52	1.55	42.2	0.34	0.89	33.1	0.53	0.94	33.4	0.38
1.75	45.1	0.53	1.80	45.5	0.34	1.37	36.3	0.37	1.25	35.7	0.37
2.10	48.4	0.77	2.09	48.8	0.58	1.43	39.1	0.53	1.44	38.6	0.50
2.29	51.0	0.67	2.26	51.2	0.56	1.68	41.7	0.50	1.64	42.3	0.50
2.63	54.7	0.62	2.75	54.8	0.78	2.00	45.4	0.38	1.96	44.9	0.34
						2.21	48.2	0.67	2.23	47.8	0.48
						2.49	51.1	0.52	2.67	51.8	0.52
						2.91	54.2	0.75	2.88	53.7	0.60
						3.14	57.0	0.75	3.15	56.9	0.71

higher value was used to determine the abscissa for that point. The curves drawn through the data were likewise traced from Fig. 2. It is apparent that thresholds measured with synchronous flashes in the two eyes are in general a little higher than the thresholds determined with alternate flashes in the two eyes. The curves, traced from the uniocular data, make a fair fit for out-of-phase thresholds, but in general, they are too low for the in-phase points, indicating that the in-phase flashes consistently yield higher critical flicker thresholds than do out-of-phase flashes or uniocular stimulation.

To simplify comparisons of results among the various conditions in the experiment, the data are plotted in terms of differences on the ordinate in Fig. 5. The top curves (A) represent the difference between flicker thresh-

olds with flashes in phase and with flashes out of phase as a function of log flash luminance. Positive deviations from zero on the ordinate indicate summation. For both *S*s there is consistent superiority of thresholds with synchronous flashes as compared with alternate flashes. Over the whole log luminance range, thresholds with in-phase flashes for *TS* average 1.43

TABLE III

MEAN CRITICAL FLICKER FREQUENCY (*CFF*) AND VARIABILITY (*SD*) IN FLASH FREQUENCY PER SECOND (*F*/SEC.) AS A FUNCTION OF FLASH LUMINANCE IN LOG MILLILAMBERTS (LOG *L*) FOR EACH *S* WITH FLASHES IN PHASE AND OUT OF PHASE

Results of *TS*

Log *L*	Flashes in phase		Flashes out of phase	
	CFF	*SD*	*CFF*	*SD*
−2.79	9.0	0.34	7.7	0.56
−1.88	12.3	0.50	9.9	0.56
−0.81	15.5	0.34	15.5	0.84
−0.51	19.7	0.34	19.1	0.77
−0.25	23.3	0.52	21.9	0.56
−0.05	25.2	0.37	24.2	0.48
0.18	27.8	0.52	27.5	0.50
0.68	32.7	0.84	30.7	0.53
0.84	34.5	0.82	33.4	0.60
1.07	36.9	0.65	35.6	0.38
1.39	40.2	0.60	39.2	0.34
1.60	43.8	0.60	42.1	0.40
1.80	47.5	0.76	45.5	0.68
2.10	50.3	0.67	48.3	0.77
2.29	53.0	0.32	50.9	0.53
2.75	56.8	0.75	54.2	0.83

Results of *GT*

Log *L*	Flashes in phase		Flashes out of phase	
	CFF	*SD*	*CFF*	*SD*
−2.46	9.2	0.33	9.1	0.47
−2.38	10.2	0.39	9.7	0.32
−1.44	12.8	0.25	12.0	0.32
−0.90	14.5	0.47	13.8	0.77
−0.73	15.5	0.34	15.0	0.48
−0.42	17.1	0.30	16.2	0.40
−0.21	19.5	0.47	17.9	0.30
−0.09	22.0	0.48	20.7	0.37
0.19	25.7	0.34	24.1	0.62
0.41	29.5	0.34	27.8	0.66
0.75	31.9	0.38	30.8	0.34
0.94	34.7	0.45	33.7	0.40
1.37	37.6	0.40	36.2	0.34
1.44	41.0	0.50	39.2	0.45
1.68	42.8	0.48	41.8	0.48
2.00	46.1	0.69	44.3	0.47
2.23	48.8	0.53	48.1	0.23
2.67	52.5	0.89	50.8	0.59
2.91	55.7	0.56	55.1	0.37
3.15	58.5	0.82	56.6	0.77

F/sec. higher than thresholds with out-of-phase flashes. For *GT* the corresponding value is 1.14 *F*/sec. Assuming phase of flash has no effect on *CFF*, the points should be distributed about the horizontal line labeled zero on the ordinate, *i.e.* the average differences would be zero. For both *S*s this null hypothesis can be rejected at better than the 0.001 level (*t*-test).

The second graph (B) shows the differences between binocular thresholds with synchronous flashes and the uniocular threshold of the 'better' eye. That is, at every level of luminance, the higher of the two uniocular flicker thresholds was subtracted from the binocular threshold. Again, synchronous binocular flashes are seen as flickering at higher flash rates than uniocular flashes of the same luminance. For *TS* the mean difference

over the range of luminance levels used was 1.37 *F*/sec.; for *GT* the corresponding value is 0.89 *F*/sec. In both instances the mean of the differences is significantly greater than zero at less than the 0.001 level.

Differences between thresholds determined with binocular regard, flashes out of phase, and the higher of the corresponding uniocular flicker thresholds are shown in Graph C of Fig. 5. The differences for both *Ss* are rather erratic, but over the whole contour the averaged differences are small. For *TS* the mean difference is 0.055 *F*/sec. (not significantly different from zero), and for *GT*, −0.25 *F*/sec. which just makes the 0.05 level of significance. Results of this study do not indicate that alternate flashes to the

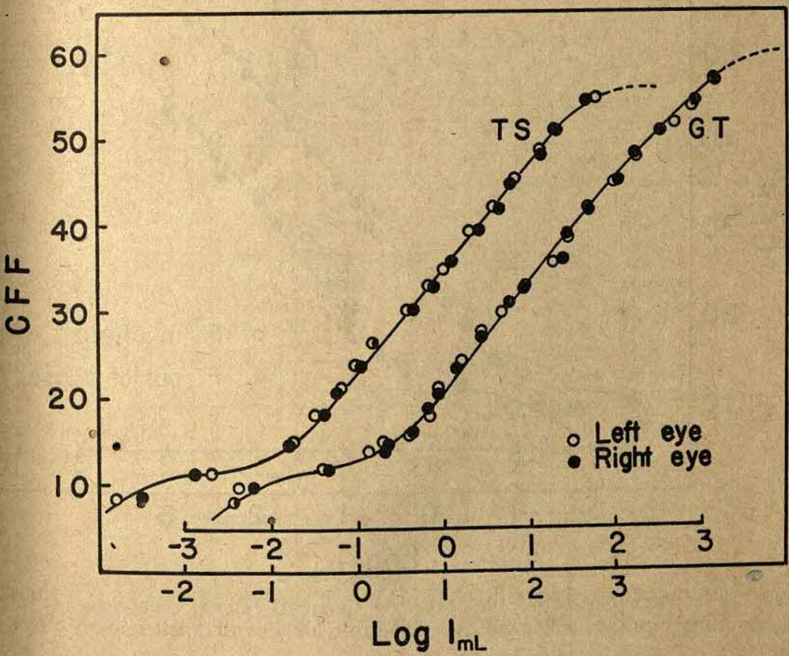


FIG. 3. FLICKER CONTOURS IN TERMS OF CRITICAL FLICKER AS A FUNCTION OF FLASH LUMINANCE

two eyes yield binocular flicker thresholds that are consistently lower than the corresponding uniocular thresholds, as has been reported.[6]

The bottom graph (D of Fig. 5) shows the differences between right and left uniocular flicker thresholds over the luminance range. Although

[6] Ireland, *op. cit.*, 282-286; Baker, *op. cit.*, 84-91.

the curve for *GT* is the more erratic, average differences are very small and insignificant for both *S*s.

O'Brien and Perrin reported that increments of thresholds with in-phase flashes over thresholds with out-of-phase flashes tend to be a constant proportion of the critical flash rate.[7] Inspection of Graph A, Fig. 5, indicates that although there is a slight tendency for the increments to increase with flash luminance (and hence with critical flash rate), the in-

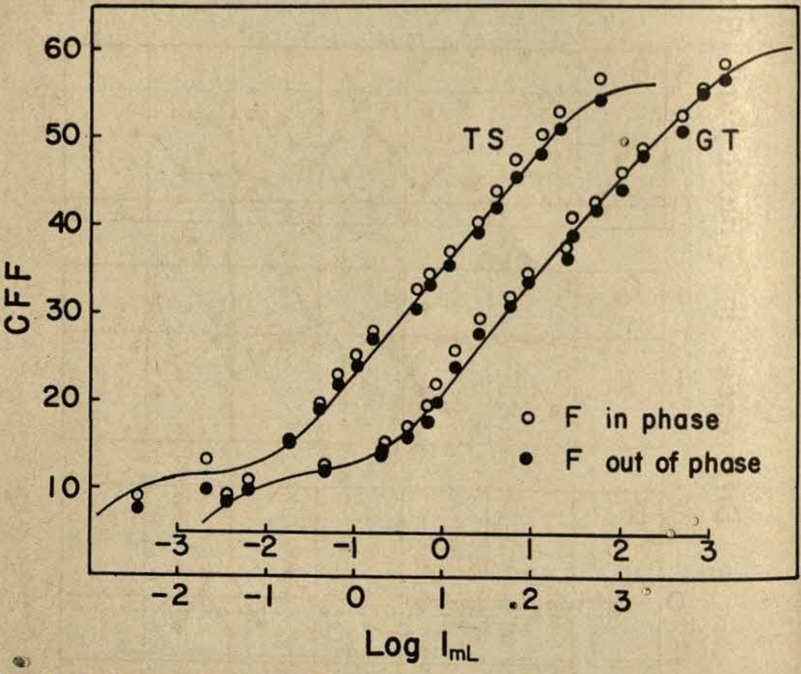


FIG. 4. FLICKER CONTOURS IN TERMS OF CRITICAL FLICKER FREQUENCY UNDER BINOCULAR REGARD AS A FUNCTION OF FLASH LUMINANCE

crease in summation is insufficient to yield a constant proportionality. The percentage increment in critical flash frequency of thresholds determined under conditions of binocular regard with flashes in phase, over the higher of the uniocular thresholds varies for one *S* from 2% to 8% and for the other *S* from about 1% to 6%. There is a slight tendency for the percentage increments to be smaller at the higher flash frequencies. The present data certainly indicate that the absolute amount of summation yielded by

[7] O'Brien and Perrin, *op. cit.*, 63.

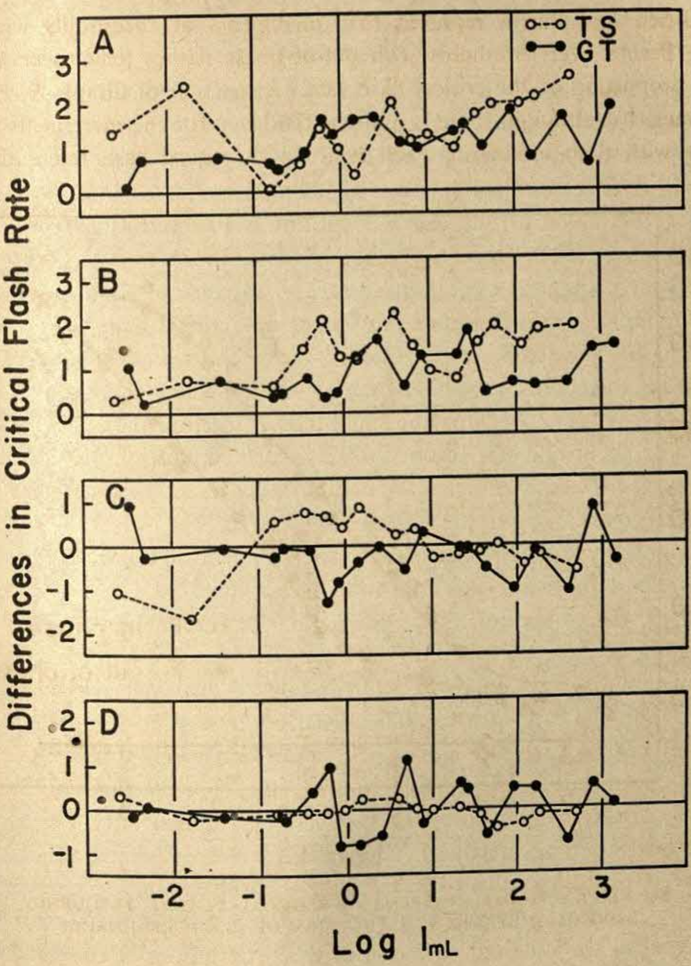


FIG. 5. COMPARISONS AMONG THE EXPERIMENTAL CONDITIONS IN TERMS OF DIFFERENCES IN CRITICAL FLICKER FREQUENCY

A = Binocular regard; thresholds with in-phase flashes minus thresholds with out-of-phase flashes as a function of flash luminance. B = Binocular thresholds with in-phase flashes minus the higher uniocular threshold. C = Binocular thresholds with out-of-phase flashes minus the higher uniocular threshold. D = Uniocular thresholds of the right eye minus the thresholds of the left eye.

in-phase flashes over the higher of the uniocular thresholds is independent of flash rate and flash luminance.

Discussion

If processes from corresponding areas of homonymous hemiretinas were so coupled in the cortex that complete, additive summation occurred, synchronous binocular stimulation should yield a visual effect identical with the visual effect of doubling the flash luminance of a stimulus projected on a given region of one retina. Doubling the stimulus-luminance results in an increase of 0.3 log unit on the abscissa. In the middle range, where the flicker contour is essentially linear, an increase of flash luminance of 0.3 log unit yields a corresponding increase in critical flash rate of roughly three flashes per second. As mentioned above, the average increment of the binocular thresholds (in-phase flashes) over the higher of the uniocular thresholds is 1.37 *F*/sec. for one *S* and 0.89 *F*/ sec. for the other, indicating that only about one-half to one-third as much summation occurs as would be expected if excitation from the homonymous hemiretinas combined additively in contributing to the unitary visual effect.

An entirely different approach to the problem of binocular interaction with regard to flicker discrimination is proposed by Crozier and Wolf. Crozier's theory postulates a population of excitable elements whose excitabilities fluctuate randomly in time. The visual effect resulting from stimulation depends upon the contribution of elements available for excitation. In the case of a constant visual effect such as *CFF*, the relationship of critical flash rate and log *I* is most adequately described in terms of the normal integral in log *I*. ('Rod' and 'cone' sections of the curve represent different but interacting populations of excitable elements and must be treated separately).[8] Thus effects of experimental variables will be most fruitfully evaluated in terms of their effects on the parameters of the fitted normal integral, viz. the maximum, the slope (inversely proportional to the standard deviation of the excitabilities of the hypothetical population of elements), and the abscissa value in log *I* at the point of inflection of the curve (the mean excitability of the population of elements).

In their comparison of binocular (flashes in phase) with uniocular flicker contours, Crozier and Wolf found, for the 'cone' segment of the curve, that the

[8] Crozier and Wolf, Theory and measurement of visual mechanisms: VIII. The form of the flicker contour, *J. Gen. Physiol.*, 25, 1941-42, 369-379; and Theory and measurement of visual mechanisms: XII. On visual duplexity, *ibid.*, 27, 1943-44, 513-528.

abscissa value of the point of inflection for conditions of binocular regard was intermediate between the corresponding values for the right and left eyes separately.[9] The asymptotic maximum (flash rate) for binocular regard was the same or higher than the maximum for the right and left eyes alone. Only slight changes in slope were observed and these were not consistent in both *S*s. That integration of the two uniocular components of excitation did in fact occur is evidenced, they argue, by a comparison of one aspect of the variability of the uniocular and binocular performances. "The law of Crozier," as it is called by Piéron,[10] states that the log variation (*PE*) of the mean flash luminance plotted against log mean flash luminance over the range of the flicker contour, approximates a linear relation with a slope of one. The variations of the points about this line, *i.e.* the *PE* of the proportionality between *PE*s and means (critical flash luminance) for the uniocular and for the binocular situations are in the proportion of $2\frac{1}{2}:1$. The authors interpreted this finding as indicating, in the binocular situation, that the 'numerical potency' of the excitable elements is doubled, although the total number of elements is only slightly or not at all increased.

The data of the experiment reported here are in several respects inadequate for conclusive comparisons with the findings of Crozier and Wolf. The data do not extend the contours over a range of flash rates and flash luminances, even for the 'cone' segment, which is sufficient for decisively fitting the normal integral to the data. The upper section of the contour, where *CFF* approaches asymptotically to a maximum, is most crucial for critically fitting a normal integral to the data.

In addition to being crucial for critical curve fitting, the shape of the extreme upper end of the flicker contour would appear to be very relevant for Crozier's theory of the flicker contour. The theory implies that as flash luminance increases, the visual system becomes saturated, *i.e.* the total population of excitable elements becomes maximally involved, and any further increase in flash luminance does not increase the threshold flash rate for detection of flicker. The flicker contour should approach asymptotically to a maximum. Other investigators find that the contour reaches a maximum and then declines at high flash luminance. Hecht and Smith reported that a surround of luminance equal to the Talbot brightness of the test-patch tends to decrease the observed decline at the upper end of the flicker contour.[11] In studies of Crozier and his associates, the test-stimulus

[9] As data from the present experiment cover only the upper end of the 'rod' segment of the F contour, consideration in discussion will be given only to the 'cone' segment.

[10] Henri Piéron, *The Sensations: Their Functions, Processes, and Mechanisms,* 1952, 329.

[11] Selig Hecht and E. L. Smith, Intermittent stimulation by light: VI. Area and relation between critical frequency and intensity, *J. Gen. Physiol.,* 19, 1936, 979-988.

does not have an illuminated surround, and yet they always find the contour to level off to a plateau at high flash luminance. Crozier and Wolf suggest that determining flicker contours in terms of averaging values from 'fusion' as well as 'flicker' end points might cause the decline.[12] Thomas, however, found that flicker contours decline beyond a maximum when measurements were in terms of just detectable flicker only.[13] Difficulties in specifying accurately the stimulus-values of flash intensity and flash frequency at high levels further complicates the problem. If the 'true' flicker contour does indeed reach a maximum and then decline, the propriety of Crozier's fitting a normal integral to the data in enunciating his statistical theory of the flicker contour is called into question.

Data continuing the contours to their maxima were not obtainable in the present experiment. The procedure of determining flicker thresholds in terms of critical flash luminance with fixed flash frequency, which was necessary to equate the stimuli in the two eyes, did not yield satisfactory thresholds at stimulus-levels higher than those reported (see Tables). At high levels, with fixed flash frequency, critical flash luminance seemed indeterminant in that no consistent set of observations was obtainable. Sometimes flicker would appear at higher flash luminances and not at lower with the same flash frequency—which is what would be expected if in fact the 'true' contour reaches a maximum and then declines, but which made determination of stable thresholds difficult.

One other reservation regarding the adequacy of the data of the present experiment in respect of Crozier's theory of binocular interaction and the flicker contour concerns the precision and statistical homogeneity of the data. At each level six sets of threshold determinations were made: (1) right and (2) left eyes alone, in terms of critical flash luminance; (3) right and (4) left eyes alone in terms of critical flash frequency; (5) binocular regard, in-phase flashes; and (6) binocular regard, out-of-phase flashes. Only one or two levels could be completed during any one session, and as a result, variability of the contour as a whole contains variance related to time of day and day to day effects.

In spite of these shortcomings of the data in respect of definitive conclusions regarding Crozier's theory, suggestive results may be obtained from an analysis along the lines proposed by the theory. Fig. 6 shows the curves for right and left eyes alone for each *S*, determined in terms of

[12] Crozier and Wolf, Theory and measurement of visual mechanisms: V. Flash duration and critical intensity for response to flicker, *J. Gen. Physiol.*, 24, 1940-41, 635-654.

[13] Thomas, *op. cit.*, 646.

critical flash frequency (Curves A and B for *TS*; Curves E and F for *GT*); binocular curves, in-phase flashes (Curves C for *TS* and G for *GT*); and binocular curves, out-of-phase flashes (D for *TS* and H for *GT*) plotted on a probability grid in log *I*. Abscissa values are shown for Curve A only. Each succeeding curve has been displaced one log unit to the right. The ordinate is in terms of 100 F/F_{max}. Fitting the normal integral to the data was done by a rough method of approximation as follows: the lowest value of F_{max} yielding a linear plot on the probability grid as estimated by eye

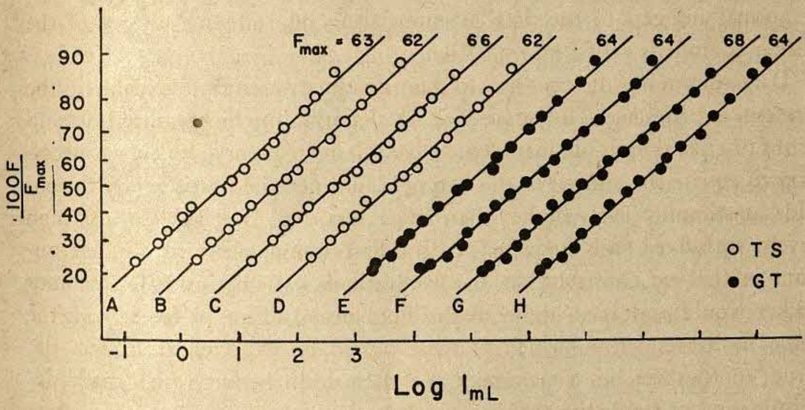


FIG. 6. CRITICAL FLICKER CONTOURS PLOTTED ON A PROBABILITY GRID
Uniocular: right eye, Curves A and E; left eye, B and F. Binocular: flashes in phase, Curves C and G; out of phase, D and H. F_{max} required for a linear plot appears at the top of every curve. Abscissa values are given for Curve A only; the remaining curves have all been displaced successively one log unit to the right.

was determined empirically. A few points at the lower ends of the contours where 'rod' effects are apparent were omitted in these plots. A straight line was then fitted to the points by the method of least squares using arbitrary linear coördinates estimated graphically. Thus approximate estimates of the three parameters of the normal integral (F_{max}, slope, and abscissa value of the point of inflection) could be determined graphically.

No changes in slope or abscissa value of the point of inflection among the four conditions of viewing showed up consistently for both *S*s. Considering the deviations of the points from the line of fit (it is apparent that these deviations are quite systematic), it is doubtful if such small and inconsistent differences in slope and abscissa value of the point of inflection that are observed would be significant. The value of the F_{max} required for linearity of the plot is, however, considerably higher for both *S*s under con-

ditions of binocular regard, flashes in phase (Curves C and G), than it is for the other three conditions of viewing. F_{max} required for approximate linearity of the plots appears at the top of each line.

In terms of the hypothesized population of excitable elements underlying visual effects, Crozier interprets two of the parameters of the fitted normal integral as follows: "The *SD* parameter [inversely proportional to the slope] is demonstrably . . . a function of the number of available neural units, while F_{max} measures the total number of elements of sensory effect producible under the conditions."[14] Thus results reported here of increased maximum and no change of slope would be interpreted as representing an increase in the frequency of contribution of the hypothetical excitable elements to the visual effect in the binocular, in-phase condition with no increase in the total number of elements contributing. Such an interpretation is plausible if the existence of cells is assumed, presumably in the cortex, on which excitation from corresponding areas in the homonymous hemiretinas converge (the isodynamic cells of Cajal[15]), and if it is further assumed that the excitation from either of the homonymous hemiretinas can stimulate these common cells, but that confluent excitation from both hemiretinas enhances their contribution to the final visual effect (increases their frequency of firing perhaps).

One more point regarding the relationship between the data of the present study and Crozier's theory merits discussion. It will be recalled from the previous discussion that one of the chief findings of Crozier and Wolf in their comparison of uniocular and binocular flicker contours concerned the variability of the proportionality between *PE* and the mean critical flash intensity (log *PE*/log I_m) at different flash rates along the flicker contour. The variability of this coefficient of variation in the two conditions (uniocular and binocular) is in the proportion of $2^{1/2}$: 1. A similar comparison of the effects of the experimental variables on the variability of response in the present experiment is not possible because the contours were determined in terms of critical flash rate rather than in terms of critical flash luminance. Uniocular contours were determined in terms of critical flash luminance, however, in order to equate the right and left stimuli (see Fig. 2), and in these instances the results are congruent with the 'law of Crozier' in that log variation (*PE*) plotted against log mean flash luminance ap-

[14] Crozier and Wolf, Theory and measurement of visual mechanisms: X. Modification of the flicker response contour and the significance of the avian pecten, *J. Gen. Physiol.*, 27, 1943-44, 289-313.

[15] Raymon y Cajal, *Histologie du Système Nerveaux de l'Homme et des Vertébrés*, 1911, II, 880.

proximates a linear relationship with slope one. When contours are determined in terms of critical flash frequency (fixed flash luminance) Crozier and his associates find the variability of the settings as a function of log I to be a complex relationship with a maximum in the middle range.[16] The data of the present experiment show no such relationship. The standard deviations of critical flash frequency are essentially independent of flash luminance (and mean flash rate) over the range of values used, *i.e.,* the sigmas, plotted as a function of mean critical flash frequency, essentially describes a straight line with slope zero. Average standard deviations for the various experimental conditions fall between 0.41 and 0.58 F/sec. There are no significant differences in size of the average standard deviations among the four experimental conditions.

The results from the present experiment and our earlier one[17] suggest there is a partial summation of excitation from homonymous hemiretinas in terms of visual effect. What rôle, if any, such summation plays in visual perception is unknown. For the most part one would suspect each of the homonymous hemiretinas has equal access to and largely equivalent effects on higher integrative centers (and ultimately motor centers) because one, the other, or both eyes yield exceedingly equivalent visual effects when stimuli of identical quality, intensity, and timing impinge upon exactly corresponding retinal areas. Yet the visual system is between 1% and 8% more effective on the average in detecting flicker when the intermittent light is flashed synchronously in the two eyes than it is with alternate flashes or with either eye alone. Perhaps 'summation' of excitation arising from corresponding retinal areas is not in fact the problem. Fry and Bartley have argued that binocular interaction in respect of brightness (Fechner's paradox) is based on effects of contour process,[18] and Werner has demonstrated central interaction of contour processes, components of which consist of excitation arising from separate eyes.[19]

If the increment of sensitivity to flicker obtaining when flashes are delivered synchronously to the two eyes could be shown to derive from or to depend upon contour-forming processes in the cortical cyclopean retina, it

---

[16] W. J. Crozier, Ernst Wolf, and G. Zerrahn-Wolf, On critical frequency and critical illumination for response to flickered light, *J. Gen. Physiol.,* 20, 1936-37, 211-228; Critical illumination and flicker frequency in related fishes, *ibid.,* 21, 1937-38, 17-56; and Intensity and critical frequency for visual flicker, *idem.,* 203-221.

[17] Thomas, *op. cit.,* 632-646.

[18] G. A. Fry and S. H. Bartley, The brilliance of an object seen binocularly, *Amer. J. Ophthal.,* 16, 1933, 687-693.

[19] Heinz Werner, Studies in contour: Strobostereoscopic phenomena, this JOURNAL, 53, 1940, 418-422.

suggests the hypothesis that the increased contribution of the hypothetical 'isodynamic' cells provides 'feedback' controlling fusional eye movements. Ludvigh believes the fusional eye movements are involuntary and controlled by visual feedback, *i.e.* the eyes orient themselves in such a way as to minimize *phenomenal* diplopia.[20] The hypothesis proposed here, on the contrary, says that the mechanism underlying the minimized diplopia depends on the orientation of the eyes so as to maximize the excitation from 'isodynamic' cells which are involved in and excited by contour forming patterns in the cyclopean retina, arising from stimulation on corresponding retinal areas. The fusional reflexes stabilize the eyes in positions which result in maximal output from isodynamic cells which in turn signifies maximal congruence of contour processes in the cyclopean cortical retina and which also provide a slight increment of visual sensitivity. Sherrington, in discussing the implications of his original experiment on binocular *CFF*, suggested that the functonal significance of the anatomical arrangement of the neuro-optic system (representation of half of each retina in the same hemisphere of the brain) was not the basis of fusion of the two uniocular images but did provide the basis of conjugate and fusional eye movements.[21] A number of phenomena of binocular visual perception could be cited lending plausibility to this hypothesis, but further speculation is gratuitous until it can be conducted within a context of more crucial experimental results.

SUMMARY

Experiments are reported comparing *CFF* measured under three conditions: uniocular regard (right and left eyes separately); binocular regard, flashes in phase; and binocular regard, flashes out of phase. At each level the right and left stimuli were equated by determining critical flash luminance for each eye separately with fixed flash frequency. Then right and left uniocular thresholds and two binocular thresholds (synchronous and alternate flashes) were determined in terms of critical flash frequency. Stimuli were of 6° subtense, centrally fixated, and fell on corresponding retinal areas when viewed binocularly.

Results show that over a range of flash luminance of about six log units (millilamberts) binocular *CFF* with synchronous flashes in the two eyes is significantly higher than binocular *CFF* with out-of-phase flashes or than uniocular *CFF*. The increment is small, being only about one half to one

[20] Elek Ludvigh, Control of ocular movements and visual interpretation of environment, *Arch. Ophthal.*, 48, 1952, 442-448.

[21] Sherrington, *The Integrative Action of the Nervous System*, 1906, 385.

third as much as would result from doubling the flash luminance of a stimulus viewed with one eye. The magnitude of the increment of summation is independent of flash luminance and flash frequency over the range of values investigated.

Although the data are inadequate in certain respects for definitive conclusions, analysis of the flicker contours along lines proposed by Crozier and his associates was undertaken. The results of the analysis tentatively support the conclusion that the increment in sensitivity to flicker under the condition of binocular regard, flashes in phase, over sensitivity obtaining under the other experimental conditions, results from enhanced contribution to the visual effect of the hypothetical population of excitable elements rather than from an increase in total number of excitable elements contributing. No significant differences in variability of performance were obtained between binocular and uniocular conditions.

A tentative hypothesis concerning the functional significance of the observed binocular summation for feedback control of fusional eye movements is suggested.

STEREOGRAM DECENTRATION AND STEREO-BASE AS FACTORS INFLUENCING THE APPARENT SIZE OF STEREOSCOPIC PICTURES

By Oscar S. Adams, Emory University

It has long been known that a change in the separation of two stereoscopic pictures will produce a change in the apparent size and distance of the objects being viewed.[1] This change, which is consistent and easily obtained, depends upon the position of the observer's eyes. Consider, for example, the situation in which the views of the right and left eyes are polarized, projected upon a screen, and then observed through polaroid spectacles. If the right and left views are superimposed on the screen, the fused picture appears in the plane of the projection-screen. If, however, the two views are so placed that the view of the right eye is to the left of the view of the left eye, *i.e.* the views are crossed, then the fused picture appears in a plane in front of the projection-screen and is reduced in size. Again, if the two views are decentered in an uncrossed relation, *i.e.* are so separated that the view of the right eye is to the right of the view of the left eye, then the fused picture appears in a plane behind the projection-screen and is increased in size. In all cases the amount of convergence required to fuse a pair of projected stereograms is controlled by varying the distance between the polarized views while holding the viewing distance of the observer constant.

In 1937 Hermans described five demonstrations involving the use of free crossed stereoscopy in which there was a decrease in the apparent size of objects with increased convergence.[2] The results of these observations together with others involving projection of the after-image in darkness and size-constancy judgments under conditions of monocular and binocular observations led him to conclude

* Accepted for publication March 24, 1954. This paper is based on a portion of a dissertation presented in partial fulfillment of the requirements for the Ph.D. degree at The Ohio State University.

[1] Wheatstone was one of the first to demonstrate the change with his stereoscope. If the mirrors of the instrument are rotated about their axes they will cause the eyes to change convergence without any change in the size of the retinal image. As convergence is increased there is a noticeable diminution in the apparent size and distance of the fused image. See Herman von Helmholtz, *Physiological Optics*, (trans. by J. P. C. Southall), 3, 1925, 312-313.

[2] T. B. Hermans, Visual size constancy as a function of convergence, *J. Exper. Psychol.*, 21, 1937, 145-161.

that convergence is a primary determiner of size-constancy. Later, Taylor demonstrated, in an ingenious experiment, that an after-image induced in total darkness would increase in size as the eyes diverged.[3] There still remains, however, the need to measure the changes in apparent size which result from decentering the stereoscopic pictures.

Another variable which is known to affect the perception of stereoscopic pictures is binocular disparity. Under appropriate conditions, fusion of disparate views of the same stimulus-object will produce an appearance of depth. The amount of disparity is normally a function of the distance that the object is from the observer, since the distance between the eyes remains constant. Even so, it is possible to produce variations in disparity while the distance between the observer and the stimulus-object remains constant. This is accomplished by presenting views obtained from positions which lie within or beyond the normal interpupillary distance. The latter case is the principle used in the construction and operation of the telestereoscope which produces views from a wide stereo-base.[4] The resulting perception is one of an exaggerated depth effect, especially for the more distant objects.

The present study is divided into two experiments. Experiment I is concerned with the effects of decentering stereoscopic pictures on the apparent size of objects appearing in the pictures. Experiment II is an extension of Experiment I, but in addition, is concerned with the joint effects of changing the size of the stereo-base at which the pictures are taken, and then changing the amount of decentration under which the stereograms are viewed.

EXPERIMENT I

Apparatus. The apparatus for Experiment I consisted of a twin projector and a Bausch and Lomb Clason Acuity Meter. The projector, designed for stereoscopic use, contained two lamp housings, two condensors, and two projecting lens having a focal length of 6.5 in. Polarized glass was placed in front of each condensor and was so oriented that the right and left beams were parallel respectively to the right and left eyepieces of *O*'s polaroid spectacles. By turning a rack and pinion gear on the projector, the distance between the two lens is decreased or increased causing the projected pictures to move closer together or farther apart.

The standard stimulus-objects consisted of four pairs of Ektachrome transpar-

---

[3] F. V. Taylor, Change in size of the after-image induced in total darkness, *J. Exper. Psychol.*, 29, 1941, 75-80.

[4] For a thorough treatment of the telestereoscope the reader is referred to Helmholtz, *op. cit.*, 3, 310-312, 343-344, 352-353, 652-688.

encies of a white target 30 cm. square. The target was photographed from distances of 6, 12, 24, and 48 m. The general surround was a level green grass expanse with a row of trees on the horizon. The photographs were made with a Rolliflex camera (75-mm. focal length) and with a stereo-base of 64 mm. The depth of field was so great that exposures could be made without changing the focus of the camera. A fifth pair of transparent stereograms was used as a control for the effect of field conditions. It consisted of a plain polarized white square projected without a figured surround. The area of this white square equaled that of the target photographed at 6 m.

The variable stimulus-object was a plain white square projected by the Clason. The standards and variable were matched for apparent brightness.

Experimental room. The stereo-projector was located 18 ft. (5.486 m.) in front of an aluminum painted screen, 6 ft. square. When projected onto the screen from this distance, the photographs of the 30-cm. white square taken at distances of 6, 12, 24, and 48 m. produced pictures of the square that were, respectively, 126, 63, 31.5, and 15.75 mm. square. Located immediately to the left of this screen was a smaller screen, 3 ft. square, on which the variable stimulus-object was projected by the Clason. This smaller screen was so placed that the distance between *O* and its center was equal to the distance between *O* and the center of the larger screen. *O* sat about 8 ft. (2.489 m.) in front of the larger screen, and between it and the projector. At this distance the visual angle subtended by the projected stimulus-object from *O*'s position was the same as that subtended by the real target from the position of the camera.

Observers. Four *O*s participated in the experiment. All of them were trained in judgments of size.

Instructions. Every *O* was instructed to judge 'apparent angular size.' This type of judgment had been used by the *E* in previous investigations, and it appeared to be superior to judgments of 'apparent absolute size.' The following instructions were given orally:

> This is an experiment in which you are to judge the size of an object which appears in a stereoscopic scene. You will notice that the scene consists of a square white target that has been photographed at different distances from the camera. [At this point the four pairs of slides were shown to the *O*.] Over to the left of the large screen you can see a smaller screen on which a plain white square is being projected. I can vary the size of this square from very small to very large. [Transition in size was demonstrated.] What I wish you to do is to tell me which direction, larger or smaller, I should make this variable square to subtend the same visual angle as the image of the standard square in the stereoscopic scene. I do not wish you to judge the actual size of the real target that was photographed, nor to try to judge the actual size of the projection on the screen. You will notice that the standard square will sometimes be in front of the variable square and at other times it will be behind it. In either case, try to make the variable square subtend the same visual angle as the standard one. You will be given as much time as you wish to make your match. I shall vary the size according to your directions. Are there any questions?

Procedure. Every *O* made 8 judgments (4 increasing and 4 decreasing in size) on each of the 4 targets at each of 13 decentrations (from 60 mm. crossed, through zero to 60 mm. uncrossed in 10-mm. steps). The pictures were presented in descending order of size as taken from 6, 12, 24, and 48 m., but the order of

decentering was varied randomly between sessions. During any one session the amount of decentration was the same for all targets, and *O* would make two judgments (one increasing and one decreasing) on each before being presented with the next one. The series of four targets was then repeated three times; hence each *O* completed eight judgments on each target at one amount of decentration (a total of 32 judgments for all targets) at each sitting. Prior to the test-matches in each session two practice matches were made of the 6-m. target, but were not recorded.

In addition to the matches made of the targets in the stereoscopic scene, each *O* was instructed to match the apparent angular size of the plain white polarized square in the same manner he had matched the target in the stereo-scene. For this series of judgments the 13 decentrations were presented in each of two sessions. The number of judgments was reduced from eight to four (two increasing and two decreasing) on each decentration in order to prevent excessive fatigue.

Results. The results of Experiment I are presented in Figs. 1 and 2 and Table I. In Fig. 1 are shown the matches of apparent size by the four *O*s for the four positions of the target at each of the 13 decentrations. The notations, 6, 12, 24, and 48 m. on the right side of the figure refer to the distances between the real target and the camera, *not* between *O* and the screen. The curves show that as the amount of decentration is changed from 60 mm. crossed viewing to 60 mm. uncrossed viewing that there is a progressive increase in the apparent size of the target at each of the four photographed positions. Furthermore, the absolute increase in size as a function of decentration decreases as the target becomes smaller, *i.e.* from 126 mm. (6-m. target) to 15.75 mm. (24-m. target).

In Fig. 2 the matches have been converted to logarithms and the individual curves combined into a mean performance-curve. Also included is the curve (broken line) for the matches of the polarized plain white square which was the same projected size as the 6-m. target. The two curves are similar enough in magnitude and slope to warrant the conclusion that the surrounding structure of the stereoscopic scene had little or no effect on the size judgments. When plotted logarithmically there appears to be no significant difference in the slopes of the 6-, 12-, 24-, and 48-m. curves.

The results of an analysis of variance performed on the transformed scores are found in Table I. A significant X^2 was obtained when the raw scores were submitted to Bartlett's test for homogeneity of variance, indicating that the variance was not homogeneous. Transforming to logarithms rendered a nonsignificant X^2. The analysis includes only the 6-, 12-, and 24-m. targets since complete data for all four *O*s could not be obtained on the 48-m. target. When there was a large amount of

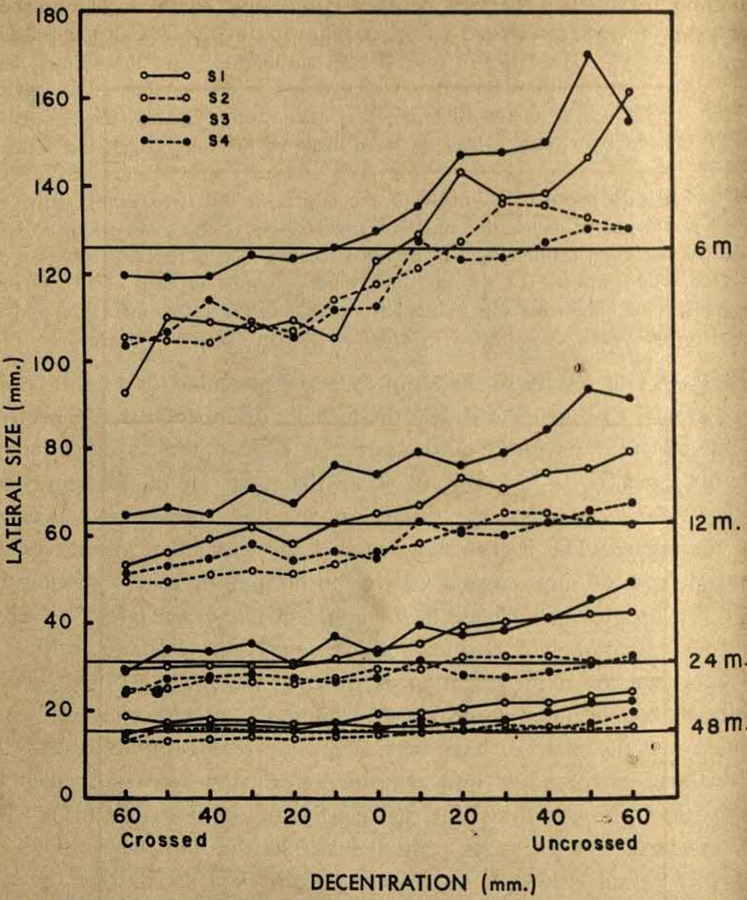


FIG. 1. MATCHES OF APPARENT SIZE BY 4 *O*s OF 4 TARGETS AT 13 DECENTRATIONS

TABLE I

SUMMARY OF THE ANALYSIS OF VARIANCE FOR EXPERIMENT I

Source	*df*	Mean Square	*F*
Decentrations (*D*)	12	.02663	30.26*
Target (*T*)	2	4.44732	5053.77*
$D \times T$	24	.00013	.15
*O*s	3	.08424	95.73*
Error	114	.00088	

* Significant beyond the 1-% level.

crossed decentration on this target the variable target could not be made small enough to satisfy the judgments of two of the *O*s. It can be observed from Table I that variation due to decentrations, targets, and *O*s were

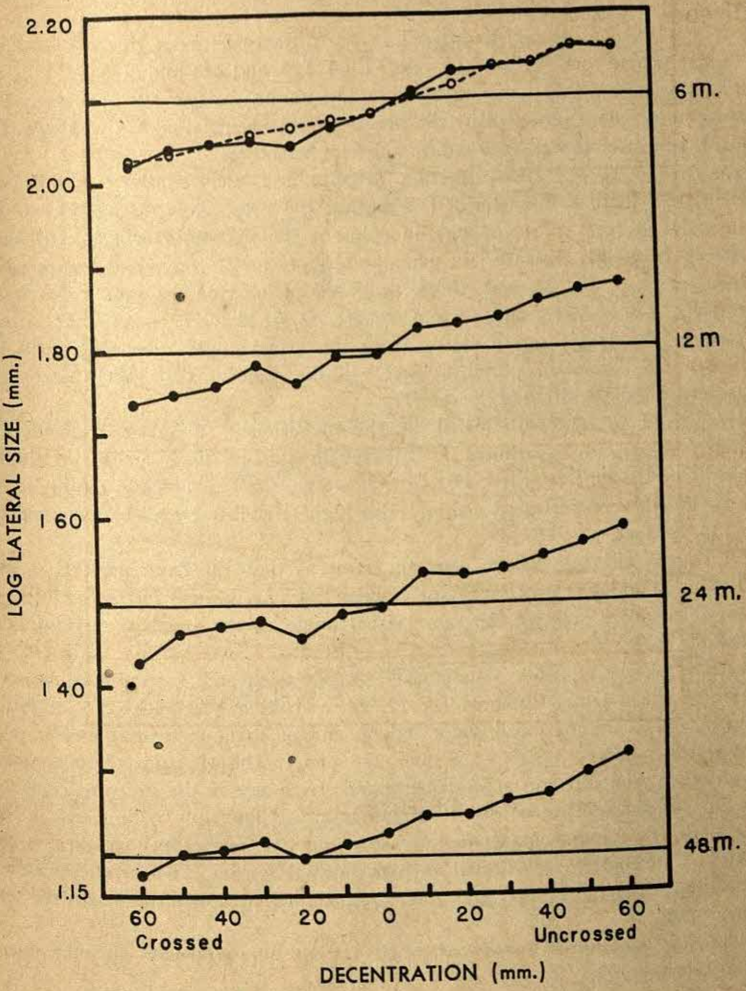


FIG. 2. MEAN MATCHES OF APPARENT SIZE

all significant beyond the 1-% level of confidence. The absence of a significant interaction between decentrations and targets is evidence that the slopes of the curves of Fig. 2 do not differ significantly from each other.

Experiment II

Apparatus. The projection apparatus was the same as used in Experiment I. The standard stimulus-objects consisted of 16 pairs of Ektachrome transparencies of a white target, 30 cm. square. As in Experiment I the target was photographed from distances of 6, 12, 24, and 48 m. In addition, at each distance four pairs of stereograms were made which differed in the distance between the two lenses of the camera. These four stereo-bases were 32, 64, 128, and 256 mm., or 1/2, 1, 2, and 4 times the normal (64 mm.) stereo-base. Again the variable stimulus-object was a plain white square projected by the Bausch and Lomb Clason Acuity Meter. The standards and variable were matched for apparent brightness.

Experimental room. The projection distance and picture sizes were slightly different from those in Experiment I. This time the stereo-projector was located approximately 18 ft. 8 in. (5.68 m.) in front of the screen. When projected onto the screen from this distance, the photographs of the 30-cm. white square taken at distances of 6, 12, 24, and 48 m. produced pictures of the square that were, respectively, 128, 64, 32, and 16 mm. square. *O* sat about 8 ft. 6 in. (2.58 m.) in front of the larger screen. Here again the visual angle subtended from *O*'s position by the projected stimulus-object was the same as that subtended by the real target from the position of the camera.

Observers. Four *O*s participated in this experiment, three of them having previously served in Experiment I. The fourth was relatively naïve in making judgments of the type required. His data, however, show remarkable conformity to those of the three experienced *O*s, and also high reliability from one experimental session to the next.

Instructions. The instructions were the same as those in Experiment I. *O* was asked to make judgments of 'apparent angular size,' *i.e.* he was instructed to make the variable square subtend the same visual angle as the standard target in the stereoscopic scene.

Procedure. Every *O* made 6 judgments (3 increasing and 3 decreasing in size) for each of the 4 target distances (6, 12, 24, and 48 m.) at each of 5 decentrations (60 and 30 mm. crossed, zero, and 30 and 60 mm. uncrossed) and at each of the 4 stereo-bases (32, 64, 128, and 256 mm.). The 16 pairs of stereograms were arranged randomly for target-distance and stereo-base in the slide-carrier of the projector, and the amount of decentration was varied randomly within and among the sessions. Two judgments (one increasing and one decreasing) on each of the 16 pairs of stereograms constituted an experimental session. Two practice matches were made on the 6-m. target (64 mm. base, zero decentration) prior to the test matches.

Under this schedule 16 sessions, each lasting approximately 40 min. were necessary for every *O*.

Results. The results of Experiment II are summarized in Fig. 3 where the mean matches of log apparent lateral size for the four *O*s are plotted against the five amounts of decentration. The two parameters in the graph are the four target positions and the four stereo-bases.

The effects of decentration and stereo-base are immediately apparent.

As the decentration is changed from 60 mm. crossed through zero to 60 mm. uncrossed, there is a progressive increase in the apparent size of

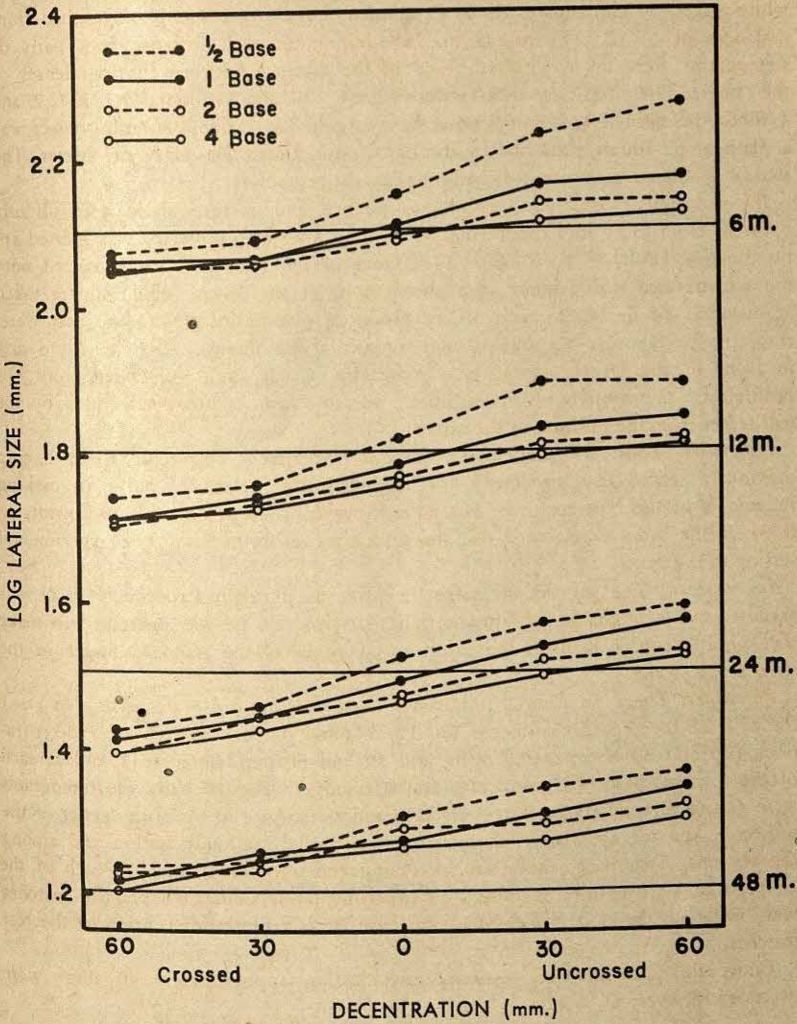


FIG. 3. MEAN MATCHES OF APPARENT SIZE OF 4 TARGETS AT 5 DECENTRATIONS AND 4 STEREO-BASES

the target at each of the four target positions. When the stereo-base of the camera is changed from half normal base (32 mm.) to four times normal base (256 mm.) there is a decrease, although not consistent for all de-

centrations, in the matches of apparent size. The effect of stereo-base appears to be greater at the uncrossed decentrations. It should be noted, however, that both decentration and stereo-base have a differential effect depending on target position, *i.e.* picture size. A non-logarithmic plot of the data in Fig. 3 would reveal this.

The mean matches of apparent size of the targets at normal base and zero decentration are as follows: 6 m. = 129.98; 12 m. = 62.14; 24 m. = 31.96; and 48 m. = 19.00. These estimates of the size of the three largest targets correspond rather closely to the actual sizes of the projected stimulus-object (128, 64, and 32 mm.), but the lateral size of the 48-m. target (16 mm.) is over-estimated by a factor of 19%. Two explanations will account for this. First, three of the points are based on data from two *O*s, and two on the basis of three *O*s. The limit of the variable target was such that *E* was unable to make a match small enough to satisfy one *O* at all. Secondly, since the size of the 48-m. target was near the limits of variable projection, there was probably a tendency toward over-estimation because of an end effect. It is concluded that if the measurements on this target were taken under improved conditions, the 48-m. curves would have lower ordinate values in Fig. 3, and the distance between the sets of curves would be constant.

TABLE II

Summary of the Analysis of Variance for Experiment II

Source	df	Mean square	F
Decentrations (*D*)	4	.15051	51.54†
Stereo-bases (*S*)	3	.04525	15.50†
Targets (*T*)	2	8.02048	6974.33†
Interactions:			
$D \times S$	12	.00292	2.54*
$D \times T$	8	.00075	.65
$S \times T$	6	.00089	.77
$D \times T \times S$	24	.00042	.36
*O*s	3	.39506	343.53†
Residual	177	.00115	

* Significant beyond the 5-% level of confidence.
† Significant beyond the 1-% level of confidence.

Table II contains the results of an analysis of variance of the transformed data for the 6-, 12-, and 24-m. targets. The 48-m. target results were not included. The *F*-ratios for decentrations, stereo-bases, targets, and *O*s are all significant beyond the 1-% level of confidence. Only the interaction between decentrations and stereo-bases is significant (P = 0.05).

The effects of change of stereo-base are more clearly shown in Fig. 4 where the mean matches of apparent size of each of the 4 targets at 60-mm. uncrossed, zero, and 60-mm. crossed decentration are plotted against the 4 stereo-base positions. The matches under the 30-mm. crossed and uncrossed decentration conditions have been omitted in an effort to increase legibility. For the three largest targets the decrease in apparent size with an increase in stereo-base varies with the amount of decentratoin. In the case of the greatest uncrossed decentration the decrease is approximately one-tenth of a log step. This changes to about one-twentieth of a log step for the most extreme crossed decentration. This interaction between stereo-base and decentration turns out to be significant at the 5-% level of confidence in the analysis of variance. Again the 48-m. results do not conform as closely as would be desired because of the smallness in size of the target and the limited number of observations obtained on it.

Discussion

The results of this study are in conformity with the observations of several investigators. It has been shown that when the stereoscopic pictures are crossed, the size of an object appears significantly smaller than when the pictures are uncrossed. The notion and basis of such a change is represented diagramatically in Fig. 5. Here the right and left eyes are represented as viewing a polarized target at crossed (c), zero (0), and uncrossed (uc) decentration. By making the geometric projections that are shown it is possible to illustrate the theoretical changes in both size and distance of the fused view as a function of the amount of decentration. The decentration controls the converging and diverging of the ocular axes, and according to Giraud-Teulon's law the size of the object is defined by the point of convergence of the axes.[5]

The theoretical distance, d, of the fused view from the O can be easily computed from the following formulas, where pd refers to the interpupillary distance, Dos to the distance between O and the screen, and dd to the amount of decentration:

$d = (pd\ Dos)/(pd + dd)$ . . . . [Crossed]
$d = (pd\ Dos)/(pd - dd)$ . . . . [Uncrossed]

The sizes which correspond to each of the distances can be determined by multiplying the various values of d by the tangent of the visual angle

[5] Henri Piéron, *The Sensations*, (trans. by M. H. Pirenne and B. C. Abott), 1952, 223.

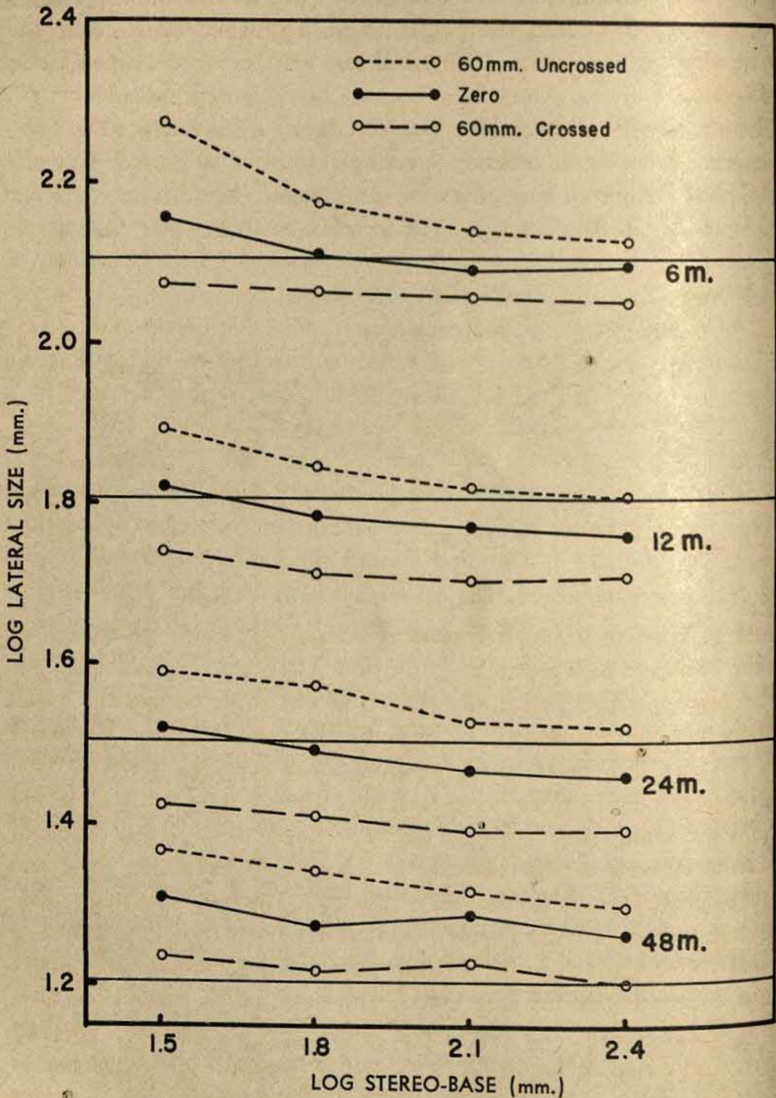


FIG. 4. MEAN APPARENT SIZE PLOTTED AGAINST STEREO-BASE WITH THE 4 TARGETS AND 3 DECENTRATIONS AS PARAMETERS

of the projected view at the distance *Dos*. Thus for the 6-m. target in Experiment I (assuming $pd = 65$ mm.) the value of d at 60-mm. crossed viewing is 1.295 m. and the lateral size is 65.5 mm. At 60 mm. uncrossed viewing d becomes 32.370 m., and the lateral size increases to 1638.0 mm.

Undoubtedly the situation is not as simple and exact as described here. When the decentration of the stereoscopic pictures is greater than the interpupillary distance the geometric projections intersect not in front of *O*, but back of the head. It must be mentioned, however, that a fair

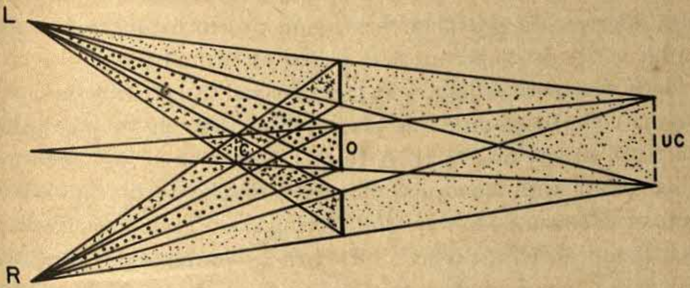


FIG. 5. GEOMETRY OF STEREO-DECENTRATION EFFECTS

amount of training is usually necessary before *O* can diverge enough to fuse the images of such pictures.

The psychophysical matches deviate from the projection of the visual angles which are represented in Figs. 1, 2, 3, and 4 by the horizontal lines for each size of target. This immediately suggests effects of size-constancy. *O*'s instructions were that he construct a square that would subtend the same visual angle in the stereoscopic scene. If no effects of size-constancy were present then the matches would have been equal and the set of curves parallel to the axis of abscissae. The results in both experiments are consistent with Thouless' notion of "phenomenal regression to the real object"[6] since almost without exception the matches made under uncrossed decentration are larger than, and those under crossed decentration are smaller than, the ones made under zero decentration. The fact that the individual curves in Fig. 1 do not all cross the horizontal line at zero decentration is taken to mean that each *O* had his own constant error in matching.

[6] R. H. Thouless, Phenomenal regression to the 'real' object, *Brit. J. Psychol.*, 21, 1931, 337-359.

The data from decentration in both experiments seem to support Hermans' emphasis on the rôle of convergence as a determiner of size-constancy. Operationally, the changes in apparent size resulted from changes in the angle of convergence. It would seem, therefore, that the data lend support to a theory which appeals to, for lack of a better term, non-visual factors as important parameters of visually perceived size. Hermans' statement to this effect is that "kinesthetic stimuli derived from the muscular activities in binocular convergence accompanying the retinal stimulation are specific cues determining size constancy."[7] If this is correct then one would predict that constancy-effects would disappear as the limits of convergence are reached. Hermans has found this to be approximately true. When his comparison-aperture was placed at 50 ft. from *O*, the standard disk which was twice as large was judged equal in apparent size at 194 ft. When the comparison-disk was placed at 400 ft., the standard disk appeared equal in size at 793 ft. A further prediction is that size-constancy should diminish with monocular observations. This is compatible with the findings of Hermans, Holway, and Boring, and Taylor and Boring.[8]

In Experiment II the effects of changes in binocular disparity are studied over a range of decentrational changes. It would be worthwhile to point out, however, that most of the disparity occurs not on the target but in the surrounding ground. Since the target was two-dimensional most of its disparity was limited to a slight change in the shape of the photographic image, *i.e.* from squareness to a suggestion of a trapezoidal shape. As distortion was more noticeable for the target photographed at the nearest positions, we would expect the effects of stereo-base to be more pronounced there. This expectation is realized when the obtained measures are plotted before transformation. When the data are transformed to logarithmic scores, however, there are no appreciable effects. This is confirmed by the analysis of variance which failed to reveal a significant interaction between stereo-base and target-distance.

On first impression it appears that the effects of stereo-base are contradictory to what would be predicted from ordinary size-distance experience. Enlarging the baseline between the two stereograms extends the stereoscopic field, but in doing so it reduces the apparent size of the objects within the field. As Helmholtz states: "Indeed when the camera

[7] Hermans, *op. cit.*, 145.

[8] A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37; D. W. Taylor and E. G. Boring, Apparent visual size as a function of distance for monocular observers, this JOURNAL, 55, 1942, 102-105.

baseline exceeds the interpupillary distance considerably, we do see reduced models that are fairly correct in form."[9] It would seem as though *O* structures his perceptual field in a manner that will bring about the best impression of orthoscopic viewing conditions. To maintain the large amount of disparity that is obtained from extreme stereo-base positions there needs to be some alteration in the size-distance relations. Inclusion of more objects having a depth-value comes at the expense of a reduction in size of the objects within the field.

There are also some practical aspects of these data. In a recent article, Smith reports the results of a questionary which was completed by 55 college students several days after they had attended a 3-Dimensional motion picture.[10] In answer to a question comparing the size of the people in three-dimensional movies with those in two-dimensional movies, 31 reported that they appeared smaller in the 3-D movie; 22 reported no difference, and two reported larger. This is perfectly understandable since the commercial three-dimensional motion pictures which use polarized light employ primarily crossed views rather than uncrossed ones. This is done to give the impression to the viewer that the scene is suspended in space between him and the screen. Uncrossed viewing would be just as effective, and probably more artistic, since it would place objects beyond the screen with about the same appearance as a stage production. The increase in apparent distance would be accompanied by an increase in apparent size as has been demonstrated in this paper.

In conclusion, it should be pointed out that unlike most methods of studying size-distance relations the one used here presents some unique and desirable features. First, the stimulus-object remains the same size and subtends the same visual angle independent of the amount or direction of decentration. Most of the phenomenal size changes can be accounted for in terms of the change in the positioning of the ocular axes. Secondly, it is unlikely that there is any appreciable change in accommodation, within limits, since none of the *O*s reported a blurring of the fused image.

SUMMARY

This paper contains the results of two experiments which investigate the effects of picture decentration and stereo-base changes on the judgments of apparent angular size of stereoscopically-viewed targets. Four *O*s

[9] Helmholtz, *op. cit.*, 675.
[10] W. M. Smith, Apparent size in stereoscopic movies, this JOURNAL, 66, 1953, 488-490.

participated in both experiments, and were instructed to make matches c apparent angular size of four targets viewed under conditions of bot crossed and uncrossed decentration. In Experiment I only the amount c decentration was varied. In Experiment II decentration was varied jointl with a change in the baseline at which the original stereograms wer photographed. Matches made of crossed views are smaller than, and thos made of uncrossed views are larger than, the matches made of the view with no decentration. Increasing the baseline of the camera between th right- and left-eye views significantly decreases the matches of apparer size. In addition, a significant interaction between stereo-base and pictur decentration is obtained. The results are interpreted as further evidenc for the importance of convergence as a parameter in the determination c apparent visual size. The effects of field factors are primarily noted wit changes in stereo-base, *i.e.* binocular disparity.

PERCEPTION OF PLACE IN A CIRCULAR FIELD

By FRED ATTNEAVE, Lackland Air Force Base

The ability of individuals to locate objects in visual space has engaged the interest of many investigators. Thus there is a voluminous literature on depth-perception, and recent studies have investigated the judgment of 'azimuth,'[1] of 'bearing,'[2] and of pointer-position on a linear scale.[3] Such studies, however, have almost invariably dealt with a single spatial dimension at a time, and the extent to which their results are generalizable to two- and three-dimensional situations is not clear. It is evident from common observation, as well as from the unidimensional studies, that place-perception is largely dependent upon certain 'anchors' or landmarks. One may reasonably suspect that as dimensionality is increased, these landmarks may acquire more important configurational properties than exist in simpler cases.

In the present study *O* was required to reproduce, from immediate memory, the locations of points presented to him singly on a circular screen in the frontal plane. The only landmark objectively present was the circular border of the field itself. The results derive their chief interest, however, from the fact that they imply the presence of certain additional landmarks of a subjective or implicit nature.

METHOD

Apparatus. The screen upon which points were presented was located 2.2 m. in front of *O*, and was 90 cm. in diameter. A translucent surface without perceptible inhomogeneity was obtained by spraying a sheet of 'plexiglass' with two coats of flat white paint. This painted plastic sheet was mounted behind a circular hole in a black wooden frame 1.5 m. high and 1.8 m. wide. The center of the screen was

* Accepted for publication March 25, 1954. From the United States Air Force Human Resources Research and Development Center. The assistance of T. P. Mitchael in the construction of apparatus for this study, and of W. H. Price and E. F. Chiburis in collecting a considerable portion of the data is gratefully acknowledged.

[1] D. C. Saltzman and W. R. Garner, Accuracy of visual estimation of azimuth position, *J. Psychol.*, 29, 1950, 453-467.

[2] S. Rogers, John Volkman, T. W. Reese, and E. L. Kauffman, Accuracy and variability of direct estimates of bearing from large display screens, *Psychol. Research Unit, Mount Holyoke College.* Memorandum Report No. 166-I-MHC 1, 1947.

[3] W. J. Carr and W. R. Garner, The maximum precision of reading fine scales, *J. Psychol.*, 34, 1952, 85-94.

1.17 m. from the floor, at approximately the eye-level of the seated *O*. Attached to the screen on *E*'s side, *i.e.* the side away from *O*, was a large removable sheet of millimeter coördinate paper. By the use of these coördinates, *E* could position a point of light on the screen with a high degree of precision. His light-source consisted of a small bulb mounted inside a freely movable cylindrical case about 7.5 cm. long and 2 cm. in diameter. At one end the case terminated in a carefully machined face 2 cm. square, in the center of which was a pinhole from which light was emitted. In presenting a point of light, *E* held the square face of the light-source against the screen, offsetting a corner of the square 1 cm. horizontally and vertically from the point at which the light was to be presented. Duration of the stimulus-light was 5 sec., controlled by an electronic timing circuit.

The only illumination in the experimental room came from two spotlights located on either side of *O* and directed at the screen. These contained 150-w. bulbs and were equipped with blue filters. Under blue illumination the orange coördinate rulings were clearly visible to *E* at a brightness low enough to prevent *O* from seeing microstructural details of the screen. The edge of the lighted area did not have the properties of a contour, since it was highly blurred, and fell upon the black frame surrounding the screen. The experimental room was dimly illuminated by light reflected from the screen, and no attempt was made to eliminate visual determinants of the horizontal and vertical from the general situation.

Directly in front of *O*, but below his line of regard, was a device by means of which he could adjust a point of light to any position on the screen he chose. This point of light was cast by a projector mounted on one end of a movable aluminum arm, the other end of which terminated in a pistol-grip handle manipulated by *O*. The arm was supported at its center of gravity by a double swivel joint on which it could be moved either horizontally or vertically; its motion was lightly braked that it would stay without further support in any position to which *O* adjusted it. A black aluminum shield, or hood, was placed over the projector to prevent stray light from shining into *O*'s eyes or upon the screen. The arm was located at a level of 84 cm. from the floor (a little higher than the bottom of the screen), but the projector at its end was mounted with an upward tilt such that the point of light was centered upon the screen when the arm was horizontal and pointed straight away from *O*. A horizontal hand-movement of about 50 cm. was necessary to move the point of light through the horizontal diameter of the screen (90 cm.); likewise, a movement of about 48 cm. was necessary to traverse the vertical diameter. The task of adjusting the light to a chosen place was easy, and preliminary tests indicated that 'motor' error was negligible: in general, *O* could superimpose his light upon a clearly defined point on the screen almost as accurately as *E* could read the setting in terms of the coördinate system, *i.e.* with an error somewhat less than 1 mm.

It was considered undesirable that *O* be allowed to maintain continuous fixation on the point at which *E*'s light appeared from the time it terminated until he made his own setting; he was, therefore, required to perform a brief task which necessitated his looking away from the screen during this interval. Mounted near his elbow (left for half the *O*s, right for the other half) was a small 'distraction-box' with two dim red lights and two push-buttons. Either the upper or the lower of these lights was automatically turned on immediately after the presentation of a

point on the screen; *O* was instructed to determine which light was on and to push the corresponding (upper or lower) button. Pushing the correct button had the effect of turning off the 'distraction-light' and turning on the point of light manipulated by *O*. If the incorrect button was pushed, a buzzer sounded. The choice of upper versus lower light was automatically made by a rotary stepping-switch, the 30 positions of which were wired in accordance with a table of random numbers.

Procedure. A single trial involved the following sequence of events. (1) *E* selected the point to be presented (see below) and positioned his light-source appropriately on the coördinate system behind the screen. (2) By pressing a button, *E* activated an electronic timer which turned his light on for a period of 5 sec. *O* carefully observed the position of this light. (3) The timing circuit simultaneously extinguished the stimulus-light and switched on one of the 'distraction-lights.' (4) *O* turned to see which of the 'distraction-lights' was on, and pressed the corresponding button; this action extinguished the 'distraction-light' and switched on the projector at the end of the aiming arm. (5) By manipulating this arm, *O* set his point of light as accurately as he could to the place on the screen at which he had just observed *E*'s light. (6) *E* marked *O*'s setting with a small dot on the coördinate paper, and connected this dot with a light line to the point at which the stimulus-light had actually been presented. (7) In preparation for the next trial, *E* pushed a button which turned off *O*'s projector.

A total of 91 different points on the screen were explored. (Their precise locations will be illustrated in the Results.) Of these points, 83 were arranged on 8 diameters with slopes of 0, ± 1, ± 3, and ∞. Half of the 64 *O*s were presented with points on the left side of the circle; the other half, with corresponding points on the right; all were given the 9 points on the vertical midline. This division, which reduced the number of settings by any particular *O* to 50, was made for reasons of experimental convenience, *i.e.* it was desirable to keep experimental sessions fairly short, and the coördinate paper used was wide enough (50 cm.) to cover only a little more than half the screen at one time. Points were presented in an irregular order, determined separately for every *O* by the sequence in which 50 shuffled plastic chips were drawn from a container. This procedure had an incidental 'randomizing' effect on the extent and direction of movement involved in settings of any particular point, since the position of the aiming arm at the beginning of a given trial was that to which *O* had set it on the previous trial.

Treatment of data. Measures of variable and constant error were computed for every position explored on the screen. For a particular position, the settings of all *O*s treated alike (*i.e.* with respect to the placement of the 'distraction-box') were transferred to a single sheet of coördinate paper in the form of a scatter-plot with the stimulus-point as origin. Quartiles and medians were then ruled off along (a) an axis corresponding to a radius of the circular screen passing through the point presented, and (b) an axis perpendicular to this radius. The interquartile ranges along these two axes were taken as measures of variable error (*VE*), and the deviations of the two medians from the stimulus-point as measures of constant error (*CE*). Since dispersions of settings were not necessarily symmetrical, two additional measures were taken, describing the distance on either dimension between the median and one of the quartiles. Results from the two groups for whom the

'distraction-box' was placed on opposite sides were combined by averaging with respect to the six measures just described.

Since the points explored were located symmetrically about both the horizontal and vertical diameters of the screen, it was possible further to condense the data into a single quadrant, as it were, by averaging symmetrical points with respect to the above measures. This process may be conceptualized most clearly as follows. Suppose that the stimulus-points and the responses of *O*s are plotted in their proper relationships on a circular sheet of paper representing the screen. This sheet is then folded twice: the bottom is folded over upon the top, and the left side is folded over upon the right. Every stimulus-point is now congruent with certain others, and response-measures pertaining to such congruent points may be averaged.

RESULTS

Fig. 1 shows the results for the entire screen as summarized in one quadrant. Places at which points were presented are indicated by dots. The quartiles of a distribution of responses are drawn as sides of a rectangle, and medians are indicated by lines crossing within a rectangle.

Variable error. Although Fig. 1 contains nearly all the information subsequently to be discussed, some of this information is more readily apprehended when presented in the form of functional relationships. Fig. 2 shows one of the two components of *VE, i.e.* interquartile range (*IQR*) along the radius, plotted against distance of the stimulus-point from the center of the screen. The value for the center point itself, which obviously constitutes a special case, was arbitrarily obtained by taking the mean of horizontal and vertical *IQR*. As one would expect from almost any version of the Weber principle, variability steadily decreases in the periphery of the circle as the well-defined edge (45 cm. from center) is approached. Error also reaches a minimum at the center, and it appears that this center-effect extends for a considerable way outward, since the region of greatest error is about a third or a quarter of the way from center to edge.

When separate functions were plotted for all five radii, it was found that the curves for the horizontal and vertical were closely similar to each other in form; so were the functions for radii 18° and 72° from the vertical (*i.e.* with slopes of 3 and of 1/3). These similar functions have been combined in Fig. 2. There is probably a genuine, though slight, interaction between distance and direction from center: functions for the five radii showed an almost complete inversion of order, or crossover, between 10 cm. and 40 cm. Whatever effect direction may have here would seem to be entirely by way of interaction, however, since the *IQR* along

the radius, averaged for all distances from center as in Fig. 3, appears very nearly constant from one radius to another.

Let us turn now to the orthogonal component of *VE*, which is measured by *IQR* across, or on an axis perpendicular to, the radius passing through the point presented. Mean values for this measure are plotted as

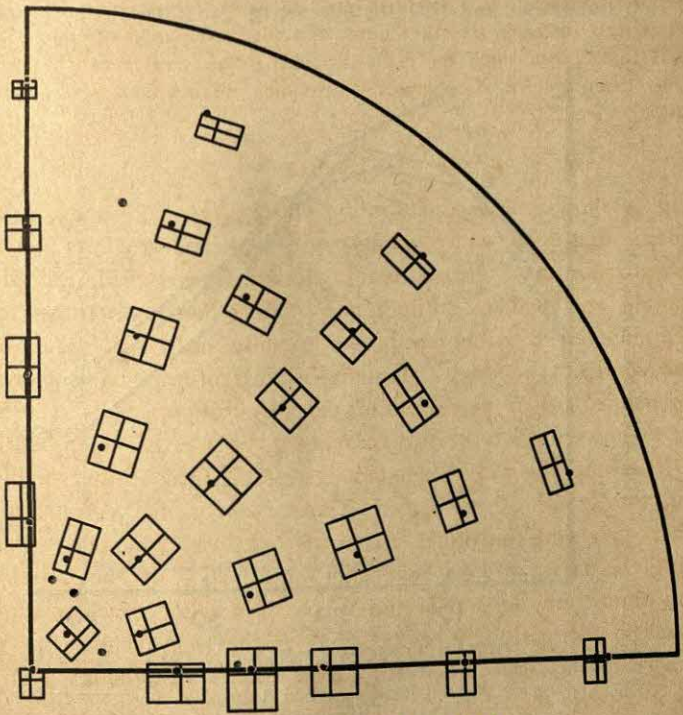


FIG. 1. RESULTS FROM THE ENTIRE SCREEN, AVERAGED AND SUMMARIZED IN ONE QUADRANT

(Dots indicate stimulus-points; rectangles show the quartiles of response-distributions along two dimensions; and points at which the lines within the rectangles cross mark the medians.)

a function of distance from center in Fig. 4. The center is arbitrarily given the same value as in Fig. 2. The other points were obtained by drawing separate functions for the five radii and averaging these at 5 cm. intervals. Although the curves averaged were not highly homogeneous in appearance (cf. Fig. 1), all five showed an initial rise in the 10–15 cm.

region, quite apart from the increase between 0 and 10 cm. which is dependent on the inclusion of the center-point.

In Fig. 5 it may be seen that cross-radius variability is less for vertical and horizontal radii than for adjacent ones. This finding, along with aspects of Figs. 2 and 4 pertaining to the central portion of the screen, is consistent with the hypothesis that *O* subjectively divided the circular field into quadrants and that the subjective 'cross-hairs' thus imposed on

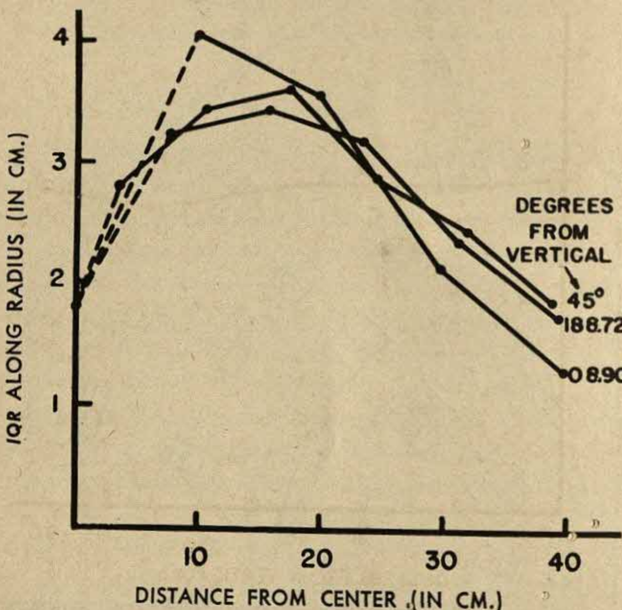


FIG. 2. INTERQUARTILE RANGE ALONG RADIUS AS A FUNCTION OF DISTANCE FROM CENTER
(Values are averaged for radii of inverse slopes.)

the screen increased the precision of settings in their vicinity somewhat as if they had been physically present. When queried, *O*s fairly consistently reported the use of some such framework.

The effect just noted, which is common to both the vertical and horizontal radii, appears to be superimposed upon a general tendency for cross-radius error to increase as the radius varies in direction from vertical to horizontal (Fig. 5). One might suppose this trend to indicate simply that the vertical component of *VE* is uniformly greater than the horizontal component, but such an interpretation would also demand a

decreasing function in the case of Fig. 3, instead of the lack of relationship actually found. An alternative hypothesis is that the implicit vertical 'cross-hair' was subjectively the more distinct or definite of the two, and accordingly that *O* was better able to judge the distance of a point from the vertical midline than from the horizontal.

Constant error. Two different types of *CE* may be described: (a) *CE* with respect to an external, gravitationally oriented frame of reference

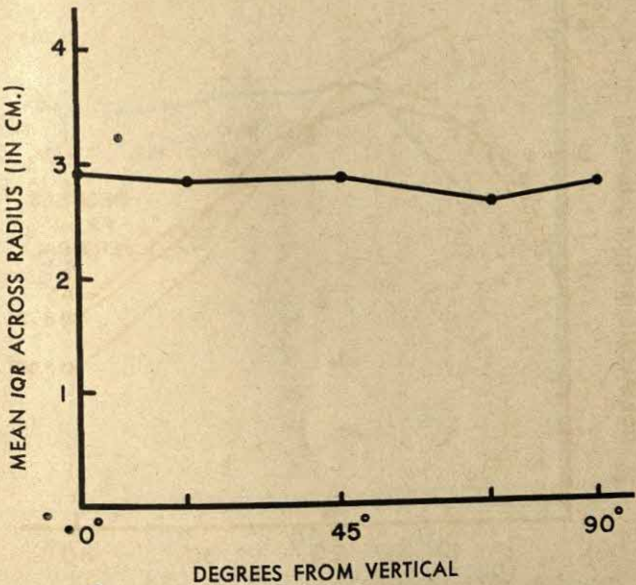


FIG. 3. INTERQUARTILE RANGE ALONG RADIUS AS A FUNCTION OF DIRECTION OF RADIUS

(*i.e.* constant up-down and left-right error); and (b) *CE* with respect to the formal characteristics of the circle (*i.e.* central-peripheral, and toward vertical diameter vs. toward horizontal diameter).

The former was calculated by taking an algebraic mean of horizontal errors, and of vertical errors, over all *O*s and all positions on the screen. The average setting was thus found to be 0.39 mm. to the left of the correct position, and 4.10 mm. below it. The reason for this considerable downward error is not known.

The second type of *CE* was essentially that which remained after the folding-and-averaging process earlier described. This process tended to cancel out constant up-down and left-right errors, since downward errors

for points below the middle of the circle became upward errors when the lower half was 'folded over' upon the top, and likewise leftward errors became rightward when the left side was 'folded over' upon the right. The question of how to treat those points which were presented directly on the vertical and horizontal diameters was settled somewhat arbitrarily, as follows. In the case of points on the horizontal midline,

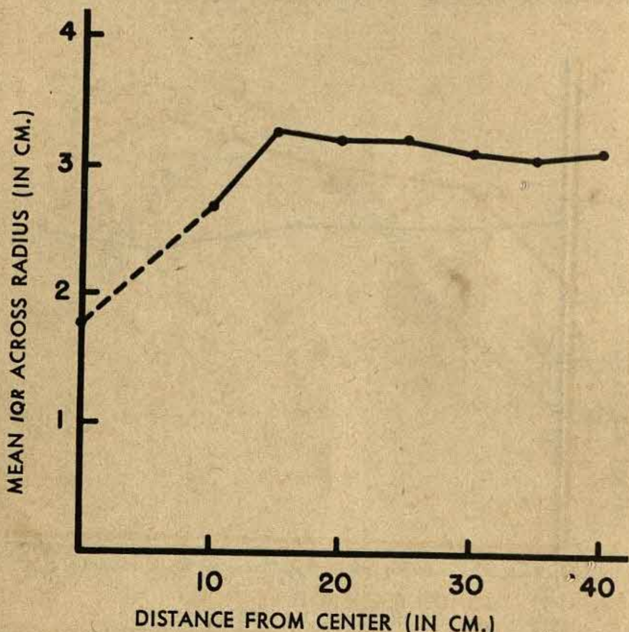


FIG. 4. INTERQUARTILE RANGE ACROSS RADIUS AS A FUNCTION OF DISTANCE FROM CENTER

only left-right *CE* was balanced out (*i.e.* by 'folding over' the left radius on the right); up-down *CE* remained unbalanced. A slightly different treatment was applied in the case of points on the vertical. It may be recalled that these points were presented to all 64 *O*s, whether the remainder of their positions were on the left or right side of the screen. Accordingly, the points on the vertical midline were treated as though they belonged to the side of the screen with which they had been associated experimentally, with the result that any genuine left-right *CE*, as well as any up-down *CE*, was balanced out. The remaining cross-radius *CE* for these points indicated to what degree *O*s tended to displace them toward the 'explored' vs. the 'nonexplored' side of the screen.

CEs based upon means (*i.e.* mean horizontal and vertical deviations) and upon medians (taken along and across radii, as described earlier) were thus averaged into Quadrant 1. The two measures of central tendency gave results so nearly identical that only *CE*s based upon medians are reported. These are drawn as vectors in Fig. 6, with calculated displacements of points (relative to the framework of the quadrant) magnified by a factor of 5 to increase legibility. It should be recalled that this information was

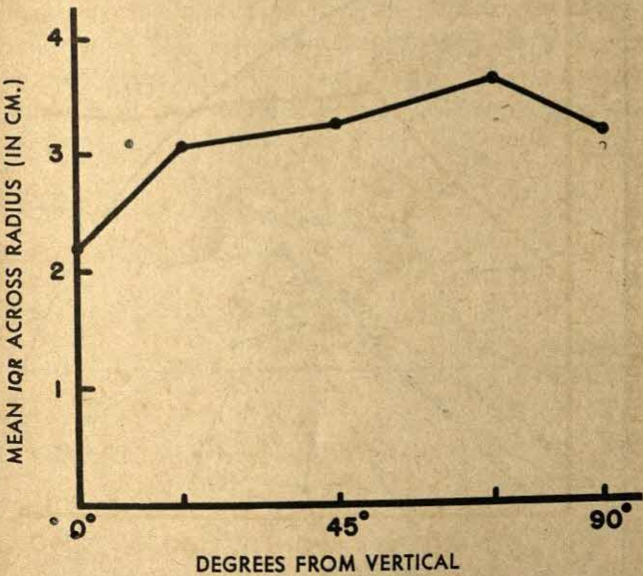


FIG. 5. INTERQUARTILE RANGE ACROSS RADIUS AS A FUNCTION OF DIRECTION OF RADIUS

contained also in Fig. 1, where the deviation of the intersection-point medians from the position of the stimulus-point represented the *CE* now under consideration.

A consistent pattern may be seen immediately in Fig. 6: *O* tended to displace any point toward the middle of the quadrant within which it appeared, *i.e.* away from the edge of the circle and also away from the vertical and horizontal diameters. This pattern of *CE* clearly confirms the hypothesis which was suggested by results on *VE*, namely, that from *O*'s point of view, the circle consisted of four fairly distinct quadrants separated by implicit 'cross-hairs.'

The points presented directly on the horizontal and vertical diameters

require special comment. In either case such points are displaced toward some central region along their respective radii, just as points elsewhere on the screen are. *CE* with respect to a gravitational framework was not balanced out in the case of points on the horizontal (see above); hence these points further exhibit the downward error which is characteristic

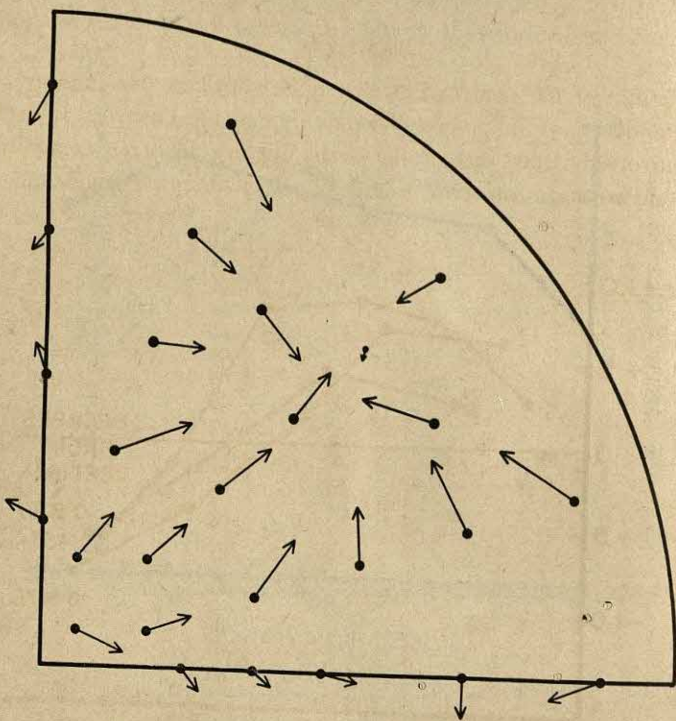


FIG. 6. CONSTANT ERRORS SHOWN AS VECTORS
(Magnitude of error is exaggerated by a factor of 5.)

of the whole screen. The leftward displacement of points on the vertical does not, however, indicate a true leftward error (see above), but instead a *CE* in the direction of the *nonexplored* side of the screen. A tentative explanation of this phenomenon will be suggested later.

Of the several functional relationships in terms of which *CE* might be described, only one shows anything that is not already obvious from Fig. 6. This is the relationship of *CE* along the radius to distance from center, shown in Fig. 7. Positive values on the ordinate indicate displacements toward the periphery; negative values displacements toward

the center. It should be noted that the direction of error reverses in a region somewhat more than half-way between center and edge, and that the points maximally displaced toward the periphery are located between 15 and 20 cm. from the center. The remarks made in connection with Fig. 2 concerning the averaging of radii apply as well to the present figure.

Discussion

The pattern of *CE* seen in Fig. 6 may be considered to indicate some 'central tendency' of judgment. What is interesting, however, is that the 'center' involved is not that of the circle, but instead of an area with no explicit demarcation on two sides, *i.e.* a quadrant. Again, one may

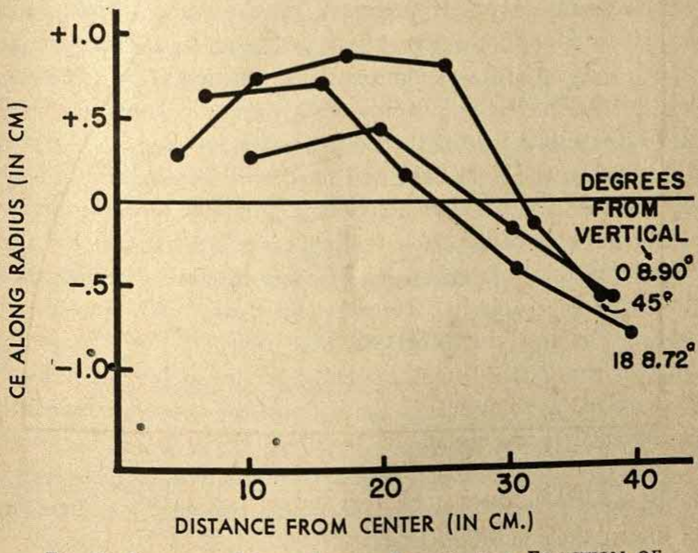


FIG. 7. CONSTANT ERROR ALONG RADIUS AS A FUNCTION OF DISTANCE FROM CENTER
(Values are averaged for radii of inverse slopes.)

suppose either that *O* was 'leveling' (*i.e.* making a point more typical or representative of the quadrant in which it appeared), or that he was 'sharpening' (*i.e.* exaggerating the deviation of a point from the nearest landmarks), but any such account must involve an assumption of the same implicit articulation of the circle. Earlier it was seen (Figs. 2, 4, and 5) that there was also a tendency for *VE* to decrease as the imaginary

'cross-hairs' of the screen were approached. It appears justifiable to infer, therefore, that subjective or implicit landmarks may have an important influence on the perception of place, and that this influence is qualitatively similar to that of landmarks objectively present.

Common observation suggests a number of forms which implicit landmarks might take. Given, for example, a pattern consisting only of two fixed dots on a vertical surface, one might expect the straight line passing through the dots, the perpendiculars to this line passing through either dot and bisecting the distance between them, the vertical and horizontal through either dot, and perhaps several other lines and curves bearing simple geometric relationships to the dots, to serve as implicit landmarks. To what extent such lines would actually have a function similar to that of the present 'cross-hairs' is a question for further investigation.

O's displacement of points presented on the vertical midline toward the nonexplored side of the screen is a matter of some intrinsic interest, even though it is an artifact from one point of view. The following is a possible explanation. If the left side of the screen is being explored, the left half of *O*'s visual field will tend to receive less stimulation from the lighted circle than the right half, and hence will become dark-adapted to a greater degree. When *O* then fixates a point on or near the vertical midline, the left side of the circle will appear somewhat brighter than the right. Assuming that brightness and area are to some extent confusable, it follows that the subjective vertical 'cross-hairs' will be shifted leftward, where the total brightness to the right of it will more nearly equal that to the left. A point presented on the objective midline will then be perceived as somewhat to the right of the 'cross-hair,' and then *O* may be expected to exaggerate its perceived deviation, just as he exaggerates objective deviations of points from landmarks elsewhere. The observable result is a displacement of points presented on the objective midline toward the nonexplored side.

Since the foregoing explanation was first considered after a pilot study in which the displacement was noted, a check upon its accuracy was possible in the present experiment. At the end of a session, *O* was instructed *verbally* to set his point upon the center of the circle. Responses showed that in most cases the subjective midline had in fact been shifted toward the explored side: of 64 *O*s, 52 erred in this direction, and only 8 in the opposite (4 responded without measurable horizontal error). Although the explanation offered is not proved by this result, its credibility is considerably enhanced.

Precisely why *O* displaced points away from nearby landmarks, objective and subjective alike, remains uncertain. Superficially it would seem that figural after-effects might have been responsible, but this hypothesis will not withstand close scrutiny. In the first place, any such effect should have influenced the point adjusted by *O* as well as the point presented, and should thus have cancelled itself out. Secondly, it is difficult to believe that a *subjective* landmark could have induced a degree of 'satiation' comparable to that induced by the black-white contour bounding the circle.

Displacement of points away from landmarks or anchors has been observed previously in unidimensional situations. For example, Carr and Garner found that in reading positions of a pointer on a linear scale, *O*s showed a *CE* toward the midpoint of a marked interval, *i.e.* away from the markers.[4] This was true, however, only for large intervals; an opposite tendency was found in the case of very small ones.

It may be that when an observed point is clearly deviant from a landmark *O* tends to emphasize and exaggerate the deviation, but that when the deviation is so small as to be difficult to detect he sometimes identifies the place of the point with that of the landmark, and hence shows an opposite error. The shape of the functions in Fig. 7 is not inconsistent with this possibility, but unfortunately no points barely deviant from the center or from the 'cross-hairs' were explored.

SUMMARY

Sixty-four *O*s (airman basic trainees) were required to reproduce, from immediate memory, the positions of single points presented on a homogeneous circular screen in the frontal plane. Half the *O*s were given points on the left side of the screen, the other half points on the right; all were presented with points on the vertical midline. A total of 91 different positions within the circle were thus explored.

It is clear that *O* utilized certain landmarks which were entirely subjective; namely, the vertical and horizontal diameters of the screen. This conclusion is based upon the fact that both *VE* and *CE* were related to distance from the implicit 'cross-hairs' much as they were related to distance from the physically present border of the circle.

The effects which both objective and subjective landmarks had on error were these:

[4] Carr and Garner, *op. cit.*, 90.

(1) Variable errors (*i.e.* the interquartile range of responses), taken along an axis oriented toward a given landmark, tended to decrease in the vicinity of the landmark. This effect was to be expected on the basis of Weber's law.

(2) Constant errors were consistently directed *away from* nearby landmarks. Accordingly, points were displaced toward the middle of the quadrant within which they were presented.

Possible generalizations of the results are suggested.

THE EFFECT OF THREE REINFORCEMENT PATTERNS ON POSITIONAL STEREOTYPES

By NORMAN R. F. MAIER and PAUL ELLEN, University of Michigan

In a previous study the effectiveness of three reward-punishment patterns in substituting a discriminative habit for a position-habit was investigated.[1] It was found that a reward-punishment ratio of 50:50 caused fewer rats to give up position-responses than ratios of either 80:20 or 20:80. The findings corresponded with an hypothesis which stated that reward-punishment ratios of 80:20 and 20:80 should produce similar results because the frequency with which a given effect (reward or punishment) will occur if a position-habit is practiced is the same for both patterns. The only difference is that in one instance reward is the predominant effect (8 out of 10 trials) and in the other, punishment is the predominant effect (8 out of 10 trials). The 50:50 pattern should be different from the other two patterns because reward and punishment occur with equal frequency. Whether or not the 50:50 pattern would be better or worse was not predicted.

The results obtained are in opposition to the predictions of reinforcement-concepts which demand very different learning effects from the uses of reward and punishment.[2] According to these concepts the reward-punishment ratio of 80:20 should produce effects very different from those of the 20:80 ratio.

Studies of abnormal fixation have shown that position-responses developed in an insoluble problem-situation are more difficult to replace with discriminative responses than are position-responses developed under standard conditions of learning.[3] To distinguish between these two kinds

* Accepted for publication December 4, 1953. This study was aided by a grant from the Horace H. Rackham School for Graduate Studies.

[1] N. R. F. Maier and Paul Ellen, Reinforcement vs. consistency of effect in habit modification, *J. Comp. Physiol. Psychol.*, 47, 1954, 364-369.

[2] C. L. Hull, *Principles of Behavior*, 1943, 124 ff; and K. W. Spence, The nature of discrimination learning in animals, *Psychol. Rev.*, 43, 1936, 427-449.

[3] Maier and Ellen, Studies of abnormal behavior in the rat: XXIV. Position habits, position stereotypes, and abortive behavior, (in press); Maier, N. M. Glaser, and J. B. Klee, III. The development of behavior fixation through frustration, *J. Exper. Psychol.*, 26, 1940, 521-546; Maier and Klee, XII. The pattern of punishment and its relation to abnormal fixations, *ibid.*, 32, 1943, 377-398; and XVII. Guidance versus trial and error in the alteration of habits and fixations, *J. Psychol.*, 19, 1945, 133-163.

of position-responses, Maier has called the first of these *position-stereotypes* and the second, *position-habits*.[4] Because the three patterns of punishment in our study with position-habits produced results contrary to expectations, it is of interest to see whether position-stereotypes will respond in a similar manner.

METHOD AND PROCEDURE

Developing position-stereotypes. The rats used as *Ss* in this study were given the same preliminary training as in the study utilizing position-habits.[5] After *S* had become adapted to the situation and had learned to jump at cards placed in the windows of a Lashley type jumping-apparatus, it was confronted with an insoluble problem. Two stimulus-cards, one with a white circle on a black background and the other with a black circle on a white background, were placed in the right and left windows of the apparatus in a pre-arranged, random sequence. On each trial one of the cards was locked in place, the sequence being random but so distributed that each stimulus-card and both the right and left windows were locked an equal number of times. The result was that no matter how *S* responded, it reached food on 5 out of the 10 trials of each day and was punished on the other 5. When *S* refused to jump within a period of 30 sec., it was tapped on the tail with a small stick. This procedure was adequate to force a choice.

Previous studies have shown that under random conditions *S* soon develops position-responses which it practices with a high degree of consistency.[6] The procedure was continued for 16 days (10 trials per day) thereby allowing *S* to practice its position-stereotype many times, receiving reward on half the trials and punishment on the other half.

Training in discriminative response. Three methods were used to replace the position-stereotype with a discriminative response and in all three the card with the white circle on the black background was made the positive-stimulus.

(1) Method A. Method A was designed to set up a reward-punishment ratio of 80:20 for practicing the previously adopted position-response and at the same time to offer consistent reward for the response to the correct card. To accomplish this, the positive card was placed on the side of *S*'s position-stereotype on 8 of the 10 trials given daily and on the side opposite the position-stereotype on the other 2 trials (the fifth and the tenth). As a result, an *S* that expressed its position-response was rewarded eight times and punished twice on each day.

(2) Method B. Method B was designed to set up a reward-punishment ratio of 50:50 for the position-response and to offer reward for all correct responses. Accordingly, the positive card was placed on the side of *S*'s position-stereotype on 5 out of the 10 trials, and on the opposite side on the other 5 trials. In this instance, an *S* practicing its position-stereotype was rewarded five times and punished five times on each day. Two variations of the 50:50 reward-punishment ratio were used. For half of the *Ss*, the negative and positive cards were on the 'position'-side

[4] Maier, *Frustration: The Study of Behavior without a Goal*, 1949, 26 f.
[5] Maier and Ellen, *op. cit., J. Comp. Physiol. Psychol.*, 47, 1954, 364-369.
[6] Maier, Glaser, and Klee, *op. cit., J. Exper. Psychol.*, 26, 1940, 521-546; Maier and Klee, *op. cit.*

on alternate trials; for the other half, the positive card was on the 'position'-side for the first five trials of a day and on the opposite side on the last five trials.

(3) Method C. Method C was like Method A, except that the positive card was on the 'position'-side on two trials per day (the fifth and the tenth) and on the opposite side on eight trials. This of course resulted in rewarding a position-response twice and punishing it eight times daily.

As in the study with position-habits, 20 days were allowed for exchanging the position-stereotype for the discriminative response (10 trials per day). An *S* was considered to have learned the discriminative-response if it made no errors on three consecutive days (30 trials). *S*s that persisted in their position-response for the full 200 trials were regarded as fixated. Only one *S* did not fall into one of these two classifications.

Subjects. Fifty-five rats were used in this study. They were taken from the stock maintained in our laboratory and were between 120 and 180 days old at the beginning of preliminary training. A split-litter technique was used to divide the *S*s into Groups A, B, and C; Groups A and B having 18 each and Group C, 19. Methods A, B, and C were used on Groups A, B, and C, respectively.

Results

As Table I shows, no striking differences among the groups A, B, and C were obtained; the procedure for developing position-stereotypes was the

TABLE I

FORMATION OF POSITION-STEREOTYPES

Group	No. Ss	Response selected		Av. No. trials response-practiced	% of total trials
		right	left		
A	18	7	11	149.5	93.4
B	18	5	13	150.5	94.1
C	19	12	7	157.6	98.3

same for all of them. In two groups, Groups A and B, *left* position-responses were in the majority; in one, Group C, *right* position-responses were favored. In all the groups combined, 43.6% of the *S*s developed right-going position-stereotypes. This result is similar to the 36.7% obtained in the previous study.[7] In both instances the *S*s were allowed to determine the position-response they would form.

The number of trials the position-stereotypes were practiced ranged from 149.5 for Group A to 157.6 trials for Group C. Since these represent from 93.4 to 98.3% of all trials given, it is apparent that the *S*s adopted a particular position-response early in the training period.

[7] Maier and Ellen, Reinforcement vs. consistency of effect in habit modification, *J. Comp. Physiol. Psychol.*, 47, 1954, 364-369.

The number of trials the position-stereotypes were practiced is only slightly less than that obtained in the study with position-habits, but the number of reinforcements given in the two studies is, of course, very different.

Table II presents the results obtained in attempting to replace the position-stereotype with a discriminative-response. Only Method B showed any degree of success in that 22.2% (4 *S*s) abandoned their position-stereotypes and 16.7% (3 *S*s) actually learned the discriminative-response. All of the *S*s in Groups A and C persisted in their position-responses throughout the training period. The results of Methods A and C combined are significantly different from the results of Method B between the 1- and 2-% levels (Chi square = 5.88).

TABLE II

INFLUENCE OF THREE REWARD-PUNISHMENT RATIOS ON POSITION-STEREOTYPES

Method	Reward-punishment ratio	No. *S*s	% persisting in position-response	% adopting discriminative problem
A	80:20	18	100.0	0.0
B	50:50	18	77.8	16.7
C	20:80	19	100.0	0.0

These results support our expectation that Method B would be unique and differ from Methods A and C. In this respect they are also like the results obtained with position-habits when these three reward-punishment ratios were used. The present results differ, however from those obtained with position-habits with respect to the direction in which Method B is unique: it is the least effective of the three methods for replacing position-habits with discriminative responses, and it is the most effective for replacing position-stereotypes with discriminative responses.

In Fig. 1 the results obtained with position-habits and position-stereotypes are presented graphically. The percentages of *S*s fixating or persisting in their position-responses are shown in relation to the three reward-punishment ratios.

In previous studies,[8] we have shown position-stereotypes and position-habits to be different in the degree to which they persisted when training methods were used to change them. We argued for a qualitative difference in persistence because bimodal distributions were obtained and had to be explained. The present results introduce another problem in that reward-punishment ratios of 80:20 and 20:80 are highly effective (relative to the

[8] See footnote 3.

50:50 ratio) in the case of position-habits, but very ineffective in the case of position-stereotypes.

Fig. 1 shows Method B to be about equally effective for position-habits and position-stereotypes. The failure to produce the degree of difference reported in previous comparisons is explained by the fact that a more

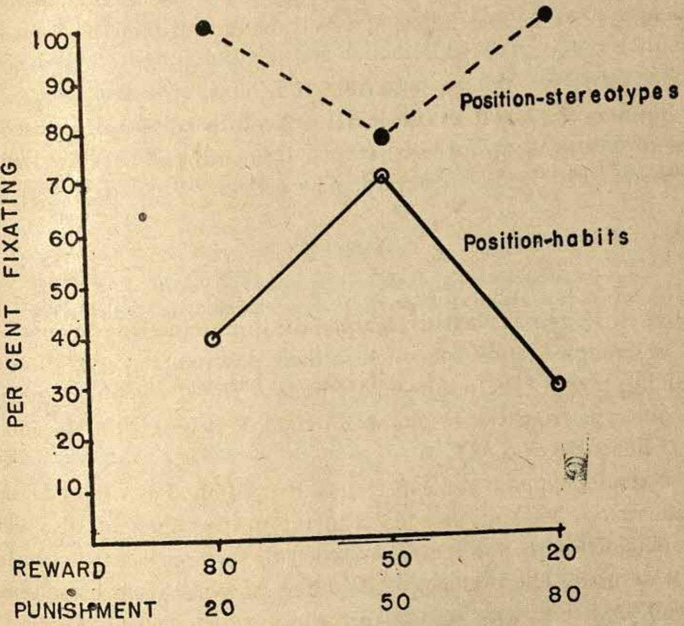


FIG. 1. PERCENTAGE OF *Ss* FIXATING THEIR POSITION-RESPONSES WHEN AN ATTEMPT IS MADE TO DEVELOP A CARD DISCRIMINATIVE RESPONSE
The upper curve (broken lines) shows the results obtained in this study when position-stereotypes were used as the initial response. The lower curve (solid lines) shows the results obtained in an earlier study with position-habits.

difficult discriminative problem was used in obtaining the results shown.[9] Increasing the difficulty of a soluble problem makes it more like an insoluble one and thus increases the number of *Ss* that fixate.

In the study of position-habits, we found Method B to be the least effective of the three when the number of *Ss* learning was used as the measure, but to be most effective, if number of trials required by the learners to adopt the new response was used as the criterion. Although all the group-differences were not significant, the trend was indicative. The

[9] Maier, *op. cit.*, 1949, 35 f; Qualitative differences in the learning of rats in a discriminating situation, *J. Comp. Psychol.*, 27, 1939, 289-331.

present study does not permit this comparison since only a few rats, those exposed to Method B, actually gave up their position-responses. The four rats that abandoned their position-responses did so after an average of 62.7 trials and the three of them that adopted the discriminative problem did so in an average of 83.3 trials. These values are very close to the corresponding scores of 58.3 and 90.0, obtained with Method B in the case of position-habits.

The data obtained on differential abortive jumping—striking positive card differently from the negative card—do, however, permit comparisons. These results are shown in Table III. The fifth column shows the percentage of *S*s in each group that jumped abortively more frequently to the negative than to the positive card. These data do not yield significant

TABLE III

FREQUENCY OF DIFFERENTIAL ABORTIVE JUMPING
(Figures in parentheses are taken from the parallel study in which position-habits rather than position-stereotypes were used.)

Method	Reward-punishment ratio	No. *S*s	*S*s showing differential abortive jumping: No.	*S*s showing differential abortive jumping: %	Av. day it appeared
A	80:20	18	11	61.1 (33.3)	9.0 (8.6)
B	50:50	18	14	77.7 (70.0)	5.3 (6.1)
C	20:80	19	15	78.9 (71.5)	5.1 (6.8)

differences, but the trends are suggestive, especially when related to those obtained with position-habits. Method A, which caused 61.1% of the *S*s to express their discrimination by this manner, is less effective than Methods B and C, which caused 77.7 and 78.9%, respectively, to jump abortively. The corresponding values obtained with position-responses in the previous study are shown in parentheses. It will be noted that all percentages obtained with position-habits are lower than those obtained for position-stereotypes but the relationship between Methods A, B, and C is the same for both studies in that Method A produces the smallest number of differential abortive jumpers and Methods B and C produce more but approximately equal numbers. The interpretation given previously is that rats under Method A have little motivation to jump abortively since the card is locked so infrequently.

The reason the frequency of abortive jumping for Method A is considerably higher in this study than in the previous one is probably due to the fact that many *S*s jumped abortively while in the insoluble problem-situa-

tion. At that stage of the training, however, differential abortive jumping was not present.

The last column of Table III shows the average number of days that elapsed before differential abortive jumping appeared. In the previous study, this was regarded as a valid measure of the rate of learning and it corresponded with the standard criterion for learning (*i.e.* the number of trials required to learn to jump to the positive card). It will be seen that with Method A, differential abortive behavior appeared after an average of 9.0 days, whereas Methods B and C required only 5.3 and 5.1 days, respectively. The difference between Methods A and B is significant at the 3-% level (with 23 *df*, $t = 2.31$). The values shown in parentheses are from the previous study with position-habits. For this measure also, the

TABLE IV

INFLUENCE OF THE FREQUENCY OF VARIATION OF POSITION OF POSITIVE AND NEGATIVE CARDS IN METHOD B

Variation	No. of Ss persisting in position-response	No. of Ss abandoning position-response	No. of Ss learning discriminative habit
1—positive card on preferred side five consecutive trials	7	2	1
2—positive card on preferred side on alternate trials	7	2	2

trend and the absolute values are very similar for the two studies. In both instances Method A is inferior and Methods B and C are close together.

Finally, we must consider the data from the two variations of Method B that were used. In one variation the positive card was placed on the side of *S*'s position-response (right or left) and the negative card on the opposite side (left or right) for the first five consecutive trials; the cards were then exchanged in the windows and left in these positions for the next five trials. Thus only one change in the position of the cards was made in 10 trials. In the other variation the position of the cards was interchanged after each jump. This resulted in nine changes. Since both Methods A and C involved three changes, these two variations of Method B constitute a control for this possible perceptual influence. Table IV shows the results obtained. As in the previous study, these variations had no effect on the basic results. For both variations, the number of *S*s persisting is seven.

For the four that abandoned, one learned with the minimal number of variations and two learned with the maximal number.

Discussion

The present study has shown that a reward-punishment ratio of 50:50 for the position-response—and consistent reward for a discriminative response—is more likely to cause rats to give up *position-stereotypes* than reward-punishment ratios of 80:20 and 20:80. In our earlier study with position-habits, the reverse was true in that the reward-punishment ratio of 50:50 was found to be the most likely of the three reward-punishment ratios to produce fixations. These data, shown graphically in Fig. 1, represent the important findings in the two experiments which require interpretation. In both studies the reward-punishment ratio of 50:50 occupies a unique position.

To account for the results, four hypotheses seem necessary. The first two of these are new; although the first one was anticipated in the experimental design, the second was not. The third and fourth hypotheses are supported by previous investigations and are not unique for this study. The suggested hypotheses are as follows.

(1) An organism becomes acquainted with its surroundings by learning the degree to which its behavior leads to certain consequences. Environments may range from orderly to disorderly ones, depending upon the degree to which 'what leads to what' is consistent.

(2) Frustration, as an intervening variable, is more likely to occur in an organism when its environment changes markedly in orderliness and conflicts with its expectation of events, pleasant or unpleasant, than when the order in a situation is consistent and expectations are fulfilled.

(3) The occurrence of frustration leads to the fixation of a highly available response.

(4) An organism with a fixated response can learn 'what leads to what,' but it cannot alter its choice because of the compulsive nature of its fixation.

In the light of these hypotheses it follows that *S*s in the experiment on position-habit, that have experienced regularity of effect (*e.g.* reward for each jump), will find the reward-punishment ratio of 50:50 to be very disturbing. The reward-punishment ratios of 80:20 and 20:80 are equally regular and should therefore be about equal to each other and less disturbing than the 50:50 ratio. That the effect (reward or punishment) is different for these two methods is not too important. As long as the irregularity is not too great, a given *S* can continue to be motivated and make wise choices. It follows then that the 80:20 and the 20:80 patterns should frustrate fewer *S*s and hence should cause more *S*s to learn than the 50:50 pattern.

Conversely, more *Ss* should fixate the highly available position-response when the reward-punishment ratio is 50:50 than when it is either 80:20 or 20:80. This conforms with the data shown graphically in Fig. 1.

Applying this same line of reasoning to the data of the position-stereotype, it follows that the reward-punishment ratio of 50:50 should be less frustrating than either of the other two methods because these *Ss* have always been in a 50:50 reward-punishment situation. The only difference is that the random 50:50 pattern has changed to an orderly 50:50 pattern. The 80:20 and the 20:80 reward-punishment patterns should be equally frustrating because these conditions represent a fairly consistent treatment for *Ss* that have become accustomed to an inconsistent pattern.

The graph shows the reward-punishment ratios of 80:20 and 20:80 to be entirely ineffectual for learning purposes since all *Ss* in both groups fixated. The fact that the total percentage of fixations in the position-stereotype groups exceeds the total percentage in the position-habit groups is to be expected because many *Ss* must have been frustrated by the insoluble problem. As pointed out elsewhere, the insoluble problem-situation causes all *Ss* to develop stereotypes and there are more fixations among position-stereotypes than among position-habits.[10] Attempts to alter position-stereotypes and position-habits of *Ss* can be a frustrating experience for them, particularly if they do not learn fast enough.

Up to this point we have considered only the manner in which the three methods used to break position-responses were frustrating. In other words, we have explained the failure to learn, rather than the nature of learning. Insofar as failure to learn is reflected in a learning score, these explanations are, however, relevant. The striking thing about the scores of the *Ss* that learned is the smallness of the differences obtained with the three methods. It is certain that as far as adopting the discriminative response is concerned, the learning rates and the percentage of *Ss* learning are unrelated, if not inversely related, measures. (The learning data with the position-habit groups showed the trials for the reward-punishment ratios of 80:20, 50:50, 20:80 to be 114.3, 90.0, and 101.6, respectively, with only the first two being significantly different. The data for the position-stereotypes fell within this range.) It seems reasonable to conclude that no strong learning variable was present in the experiment. This failure to obtain a sizeable difference in learning occurred despite the fact that the conditions of reinforcement were greatly varied.

If we assume that the learning factor is one of associating reward with

[10] Maier, *Frustration: The Study of Behavior without a Goal,* 1949, 33 ff.

one card and punishment with the other, then all learning conditions are similar in that magnitude of differences to be discriminated and the number of variables are the same for both groups and all methods.[11] All three methods permit either the learning of an association between a jump to the negative card and punishment, an association between a jump to the positive card and reward, or both. Any of these associations will be adequate for performing satisfactorily in the discriminative problem.

A more important and relevant variable seems to be that of motivation. Punishment caused *Ss* to use their learning to jump abortively in such a manner as to reduce pain. For both position-habits and position-stereotypes, the 80:20 reward-punishment pattern was the least conducive to the development of abortive jumping, whereas the other two conditions were equally conducive. Likewise, the 80:20 reward-punishment pattern was poorest when measured in terms of the number of days required for abortive jumping to occur and again the other two conditions were better and equal.

The reasons for interpreting these differences in abortive jumping as motivational rather than as learning factors are as follows.

(1) Reinforcement learning should demand that all three conditions be different and form a linear relationship, whereas a motivational interpretation permits an all-or-nothing relationship in that something is either worth doing or it is not.

(2) If learning differences occurred, they should have been reflected more clearly in the score measuring the rate of discriminative learning, but motivational differences permit performance and learning scores to differ. *Ss* can continue to jump to the negative card even after they have learned it will punish them.

(3) A learning interpretation would be difficult to reconcile with the bimodal distributions obtained by each method, whereas a motivational interpretation fits readily with the frustration-motivation dichotomy.

Since the frustration does not necessarily interfere with cognitive learning, *S* can still learn that one card punishes and the other one rewards. As a result *S* can show its anticipation of reward and punishment by jumping directly to the positive card and abortively to the negative card. It cannot, however, follow the positive card from side to side because of its fixated position-compulsion. The reason the position-response is perpetuated in both groups is because this is the response in progress at the time frustration occurs and nothing has happened to make another response more available.

---

[11] Maier, The specific processes constituting the learning function, *Psychol. Rev.*, 46, 1939, 241-252; N. R. F. Maier and T. C. Schneirla, *Principles of Animal Psychology*, 1935, 348 ff.; Mechanisms in conditioning, *Psychol. Rev.*, 49, 1942, 117-134.

If we translate the implications of this study into every day terms, we are saying that if an individual is consistently rewarded for one thing and consistently punished for another thing, he develops certain expectations. If we reversed reward and punishment, this might be disturbing, but it would not be as disturbing as inconsistent reward and punishment.

On the other hand, if an individual lived in a world where punishment and reward occurred at different times in connection with the same behavior, he would learn to expect this random condition. If the irregularity disappeared and either reward or punishment consistently occurred in connection with a particular behavior, the individual would be disturbed. He would be more disturbed by this sharp change than if the reward and punishment occurred with the same frequency as before, but followed a different plan. This implication, that the degree to which a given experience is disturbing, confusing, or frustrating, depends not only upon the situation but also upon the individual's expectations, suggests problems created by other data. The Brunswik[12] and Humphrey[13] studies of probability, the recent investigations directed toward building behavior models,[14] and the special treatment given to partial reinforcement,[15] all permit one to interpret expectation as a potential intervening variable—in addition to learning—in determining an organism's response to a given situation. It may also throw light on a previous finding which demonstrated that guidance (influencing an animal's choices by using the hand to direct it to jump to the positive stimulus-card) given during the initial learning of a discrimination prevents the later appearance of fixations when the animal is subjected to frustration.[16] It is possible that early guidance so influences expectations that the later changes in the situation are less frustrating.

---

[12] Egon Brunswik, Probability as a determiner of rat behavior, *J. Exper. Psychol.*, 25, 1939, 175-197.

[13] L. G. Humphreys, The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions, *J. Exper. Psychol.*, 25, 1939, 141-158; Acquisition and extinction of verbal expectations in a situation analogous to conditioning, *idem*, 294-301.

[14] R. R. Bush and Frederick Mosteller, A mathematical model for simple learning, *Psychol Rev.*, 58, 1951, 313-323; and W. K. Estes, Toward a statistical theory of learning, *ibid.*, 57, 1950, 94-107.

[15] John Dollard and N. E. Miller, *Personality and Psychotherapy: An Analysis in Terms of Learning, Thinking and Culture,* 1950, 16 f.; W. O. Jenkins and J. C. Stanley, Partial reinforcement: A review and critique, *Psychol. Bull.*, 47, 1950, 193-234; and H. C. Wilcoxon "Abnormal fixations" and learning, *J. Exper. Psychol.*, 44, 1952, 324-333.

[16] Maier and Ellen, Studies of abnormal behavior in the rat: XXIII. The prophylactic effects of "guidance" in reducing rigid behavior, *J. Abnorm. Soc. Psychol.*, 47, 1952, 109-116.

In previous studies we described as frustrating such things as (a) an insoluble problem, (b) pressure from behind, (c) barriers preventing escape, and (d) persistent or severe punishment.[17] We now added another; namely (e) consistency or inconsistency of results that conflict with expectations. At the same time doubt has been cast on the validity of the concept of punishment *per se* as a frustrating agent. It is rather interesting that Marquart and Arnold,[18] using only verbal 'right' and 'wrong' statements, have duplicated the basic findings of Marquart who used electric shock for wrong responses in the production of fixations in college students.[19] This too suggests that failure (which is based on level of aspiration or expectancy), not punishment *per se,* is the frustrating factor. The evidence that prolonged or intense punishment will produce frustration and the resulting fixations[20] is not, however, invalidated and is consistent with the above interpretation if we accept the view that the threshold of frustration must be passed to obtain its effects.[21]

Pronounced differences among individuals occur in certain situations because whether any or all of the factors mentioned above frustrate depends upon a given individual's threshold. Using group-scores without examining the nature of the distribution hides the crucial distinction. Problems or confronting an individual with obstacles need not be frustrating. A problem-situation may stimulate problem-solving behavior and if the situation is challenging, it may be highly motivating. Experimenters who claim that frustration increases efficiency use the term *frustration* to mean that an organism has been confronted with an obstacle.[22] A situational definition of this kind is so inclusive that opposite kinds of behavior are included and therefore confuse the issue rather than clarify the processes involved.

---

[17] See footnote 2.

[18] D. I. Marquart and L. P. Arnold, A study in the frustration of human adults, *J. Gen. Psychol.,* 47, 1952, 43-63.

[19] Marquart, The pattern of punishment and its relation to abnormal fixation in adult human subjects, *J. Gen. Psychol.,* 39, 1948, 107-144.

[20] Maier and Ellen, Studies of abnormal behavior in the rat: XXIII. The prophylactic effects of "guidance" in reducing rigid behavior, *J. Abnorm. Soc. Psychol.,* 47, 1952, 109-116; Maier and R. S. Feldman, XXII. Strength of fixation and duration of frustration, *J. Comp. Physiol. Psychol.,* 41, 1948, 348-363; Maier and Klee, XII. The pattern of punishment and its relation to abnormal fixations, *J. Exper. Psychol.,* 32, 1943, 377-398; and R. L. Solomon, L. J. Kamin, and L. C. Wynne, Traumatic avoidance learning: The outcomes of several extinction procedures with dogs, *J. Abnorm. Soc. Psychol.,* 48, 1953, 291-302.

[21] Maier and Ellen, The effects of lactose in the diet on frustration-susceptibility in rats, *J. Comp. Physiol. Psychol.,* 44, 1951, 551-556.

[22] I. L. Child and I. K. Waterhouse, Frustration and the quality of performance: II. A theoretical statement, *Psychol. Rev.,* 60, 1953, 127-139; F. McKinney, G. B. Strother, R. R. Hines, and R. A. Allee, Experimental frustration in a group-test situation, *J. Abnorm. Soc. Psychol.,* 46, 1951, 316-323.

Summary

Position-stereotypes developed in an insoluble discriminative problem-situation (reward and punishment given at random) were subjected to three methods of training, all designed to cause the *S*s (rats) to substitute a discriminative response for the response in progress. Each method gave reward for a response to the positive member of a pair of stimulus-cards and punishment for a response to the negative member, but the reward-punishment ratios for the position-response were 80:20, 50:50, and 20:80 for Methods A, B, and C, respectively. The results obtained showed that for Methods A and C, all the *S*s persisted in their position-stereotypes and demonstrated position-fixations, whereas for Method B, only 77.8% of them showed such position-fixations.

These results are in contrast to those obtained in a parallel study with position-habits, where the results of Method B also differed from those of A and C. Instead, however, of being the best for learning, it caused more fixations than either Methods A or C. To account for the results described above as well as other learning data—such as trials required to learn and the two measures for the development of differential abortive behavior—four postulates seem necessary.

(1) An organism becomes acquainted with its surroundings by learning the degree to which its behavior leads to certain consequences. Environments may range from orderly to disorderly ones, depending upon the degree to which 'what leads to what' is consistent.

(2) Frustration, as an intervening variable, is more likely to occur in an organism when its environment changes markedly in orderliness and conflicts with its expectation of events, pleasant or unpleasant, than when the order in a situation is consistent and expectations are fulfilled.

(3) The occurence of frustration leads to the fixation of a highly available response.

(4) An organism with a fixated response can learn 'what leads to what,' but it cannot alter its choice of card because of the compulsive nature of its fixation.

'ISOLATION' AND THE LAW OF EFFECT

By Leo Postman and Pauline Austin Adams,
University of California

Selective reinforcement of word-number associations has frequently been used to test Thorndike's law of effect. *S* guesses a number to each of a series of words. Typically, the large majority of the guesses is called 'wrong,' and a small proportion is reinforced by the announcement of 'right.' When the stimuli are presented again, a larger percentage of rewarded than punished responses is usually repeated. Thorndike interpreted such results as evidence for the automatic action of rewards, but his interpretation has been challenged on several grounds.[1]

Rewards may be effective not because they automatically strengthen stimulus-response connections but because they provide information which the *S* can acquire and use if and when he is motivated to repeat rewarded responses. Such an interpretation seems to be supported by the finding of Wallach and Henle that rewards fail to strengthen stimulus-response connections in the absence of a motive to learn.[2] In a typical Thorndikian situation, they informed their *Ss* that the experiment was concerned with extra-sensory perception (*ESP*) and that responses called 'right' on one trial might or might not be correct on the next trial. Under these conditions, rewarded responses were neither repeated nor recalled more often than punished ones. The results of Wallach and Henle were not, however, conclusive for two reasons. First, the stimulus-list was presented several times in succession, with different stimulus-response associations reinforced on successive trials. This procedure gives rise to an unknown, and probably considerable, amount of interserial interference. Secondly, the analysis of Wallach and Henle fails to draw a sharp distinction between the conditions of learning and performance. The initial instructions, designed to eliminate the learning motive, may influence acquisition, performance, or both.

In a previous communication we presented the results of a study in which we attempted to separate experimentally the effects of reward on learning and performance.[3] All *Ss* were given the *ESP*-instructions used by Wallach and Henle, but after one training trial were divided into groups receiving different performance instructions. The performance-instructions stated, for different groups, that previously rewarded or punished responses (a) would henceforth be correct, (b)

---

* Accepted for publication March 15, 1954. This research was facilitated by a grant from the Behavioral Sciences Division of the Ford Foundation.

[1] E. L. Thorndike, *An Experimental Study of Rewards*, 1933, 1-72.

[2] Hans Wallach and Mary Henle, An experimental analysis of the law of effect, *J. Exper. Psychol.*, 28, 1941, 340-349; A further study of the function of reward, *ibid.*, 30, 1942, 147-160.

[3] Leo Postman and P. A. Adams, Performance variables in the experimental analysis of the law of effect, this Journal, 67, 1954, 612-631.

would be incorrect, (c) might or might not be correct. Tests of repetition and recall were given under these different instructions. Both rewarded and punished responses were repeated at a better than chance level, under most of the conditions of testing, even when performance instructions did not call for such repetition. Performance-instructions did, however, significantly influence the level of repetition: the more positive the instructions the higher was the frequency of repetition. Finally, rewarded responses were in general repeated and recalled more frequently than punished ones. We were led to conclude that rewards strengthen stimulus-response associations even in the absence of a learning motive. It was equally clear, however, that performance-instructions modulate, over a wide range, the effects of reinforcement and exercise.

Our evidence for the effectiveness of reinforcement is, however, like that of Thorndike, subject to another criticism. A few rewards scattered through a long series of punishments may create conditions for the 'perceptual isolation' of the rewarded pairs. Rewarded responses may be repeated not because they are reinforced but because they are recalled better as a result of their perceptual isolation. This argument was advanced by Zirkle, who presented experimental evidence that rewarded responses are repeated more frequently than punished ones only if the rewards are isolated events in an homogeneous series of punishments.[4] Zirkle's data also showed that a gradient of repetition (spread of effect) appears only around isolated responses. At the same time, the results of his experiments clearly indicated that isolation *per se* does not lead to the repetition of a response: a combination of reward and isolation appeared to be required. Zirkle concluded that "a satisfier isolates tendencies which are at hand when it happens,"[5] *i.e.* the probability of repetition is increased by reward, but only if the reward is perceptually isolated.

Zirkle's experimental evidence is, however, open to the same criticism as the findings of Wallach and Henle. The lists of stimulus-items were presented a large number of times in rapid succession, and different stimulus-response associations were reinforced on successive trials. Again the factor of interserial interference was not controlled. It seems important, therefore, to reopen the question whether perceptual isolation is a necessary condition for the effectiveness of reward. In our earlier study we had concluded that the repetition of rewarded responses is, to a limited extent, independent of the *S*'s motivation to learn and repeat such responses. If the repetition of rewarded responses can be shown to be similarly independent of perceptual isolation, the case for the effectiveness of reinforcement in verbal learning would be further strengthened.

The following experiment was designed to evaluate the relative importance of reinforcement and 'isolation' as determinants of response-repetition. In a typical Thorndikian situation, response-repetition was related to systematic variations in the relative frequency of rewards and punishments. To obtain unequivocal evidence for the sheer effects of differential reinforcement, the experiment was conducted under conditions of incidental learning. Are rewarded responses repeated

---

[4] G. A. Zirkle, Success and failure in serial learning: I. The Thorndike effect; II. Isolation and the Thorndike effect, *J. Exper. Psychol.*, 36, 1946, 230-236; 302-315.

[5] Zirkle, *op. cit.*, 312.

more often than punished ones when *S* is not motivated to learn and regardless of the perceptual isolation of the rewards?

Method

Experimental design. The stimulus-series consisted of 21 three-letter nouns to each of which *S* responded with a number from 1 to 10. The experiment was introduced as a study of extra-sensory perception, and *S* was informed that responses called 'right' on the first trial might or might not be correct on the next trial. The instructions were exactly the same as those previously used by Postman and Adams.[6] A statement was added that a considerable number of hits might occur in a single series by chance alone.

The major experimental variable was the degree of isolation of rewarded responses. There were three experimental conditions, in which the relative frequency of rewarded and punished responses was systematically varied. In Condition I, three responses, viz. those in Positions 4, 11 and 18 were rewarded. In Condition II, 10 responses were rewarded and 11 responses were punished. The positions of the rewarded responses were randomly scattered through the series, but single alternation of rewarded and punished responses was avoided as far as possible. Finally, in Condition III, 18 responses were rewarded and three responses were punished. The punished responses were in Positions 4, 11 and 18, *i.e.* in the same positions (and associated with the same stimulus-words) as the rewarded responses in Condition I. Thus the three conditions cover a range from a high degree of isolation to a high degree of crowding of rewarded responses.

Procedure. The stimulus-words, typed in capital letters on index cards, were presented at the rate of approximately 1 every 5 sec. The experimental procedure comprised a training-series, a test-series and a recall-test.

During the training-series, differential reinforcements of 'right' and 'wrong' were administered according to the schedules outlined above. The training-series was followed by a test-series in which the stimulus-words were presented in a different order.[7] Prior to the test-series, *S* was reminded that responses called 'right' on the first trial might or might not be correct again. No announcements of 'right' or 'wrong' were made during the test-series.

Following the test-series, a recall-test was given. *S*'s task was to recall the numbers with which he had responded to each of the words during the training-series and to indicate whether these responses had been called 'right' or 'wrong.' A new order of presentation was used for the recall-test.

Recall control groups. There were three recall control groups, one for each of the experimental groups. These control groups were given a recall-test immediately after the training-series. The performance of the control groups provides a measure of recall which is independent of the practice provided by the test-series.

Subjects. There were 20 *Ss* in each of the experimental groups and in each of the recall control groups.[8] All *Ss* were undergraduate students at the University of California. They did not know the purpose of the experiment.

---

[6] Postman and Adams, *op. cit.*, 612-631.

[7] The same restrictions on the order of presentation were imposed as in the previous study of Postman and Adams, *op. cit.*, 612-631.

[8] *Ss* were discarded and replaced if they failed to follow instructions, *e.g.* gave

Results

(1) Repetition of rewarded responses. Table I shows the average number of rewarded and punished responses repeated on the test-trial under each of the experimental conditions. To facilitate comparison among the conditions, all averages were converted into percentages and are presented graphically in Fig. 1.

In evaluating the effectiveness of rewards, Thorndike assumed that with 10 numbers to choose from the chance probability of a repetition is 0.10. By this criterion, rewarded responses are repeated above chance at the better than 1-% level of confidence under all three experimental conditions. The assumed chance level has been questioned, particularly in view of the evidence for systematic guessing habits on the part of the *Ss*[9] No more adequate estimate of chance-level

TABLE I

Average Number of Rewarded and Punished Responses Repeated by the Various Experimental Groups

Condition	Rewarded responses	Punished responses
I	.75	2.45
II	1.95	1.75
III	4.15	.35

is, however, available. In view of the high levels of confidence it seems fair to conclude that rewards had significant effects under all conditions.

Whatever the error in the chance-estimate, it must be assumed to be equal for all experimental conditions. The error should, therefore, not influence the answer to the critical experimental question, viz. whether the relative frequency of repetition of rewarded responses varies as a function of perceptual isolation. Analysis of variance of the repetition scores (weighted for opportunity of occurrence) fails to yield a significant difference among conditions ($F = 0.32$). If the data are taken at face value, there is some indication of a lower level of repetition in Condition II, suggesting that rewards are least effective when they alternate with punishments in about equal proportions. Highly crowded and highly isolated rewards, however, are almost equally effective.

(2) Repetition of punished responses. According to Thorndike, the sheer occurrence of a response may overcome whatever weakening effects punishment exerts on a stimulus-response association.[10] There is some support for this view

their responses before seeing the stimulus-words. They were also discarded if they persisted in giving the same number or a very small range of numbers throughout the series. Such persistence of responses occurred most frequently under Condition III, suggesting that there may be cumulative effects of reinforcement within a series.

[9] Evidence for such guessing habits is presented by W. O. Jenkins and L. M. Cunningham, The guessing-sequence hypothesis, the 'spread of effect,' and number-guessing habits, *J. Exper. Psychol.*, 39, 1949, 158-168.

[10] For a review of the literature on the rôle of punishment see Postman, The history and present status of the law of effect, *Psychol. Bull.*, 44, 1947, 489-563.

in the present study, just as there was in our previous experiment. Applying the same criterion as before, punished responses are repeated at a better than chance level in Conditions I and II but not in Condition III. Analysis of variance of the weighted repetition scores, however, shows the difference among conditions not to be significant ($F = 0.36$). Disregarding the lack of statistical significance for the moment, it appears that punishments are least effective in weakening associations when rewards and punishments are alternated in about equal proportions.

(3) Differential effectiveness of rewards and punishments. Under all conditions of the experiment, rewarded responses are repeated relatively more frequently than punished ones (Table I and Fig. 1). Due to high variability, only one of the tests

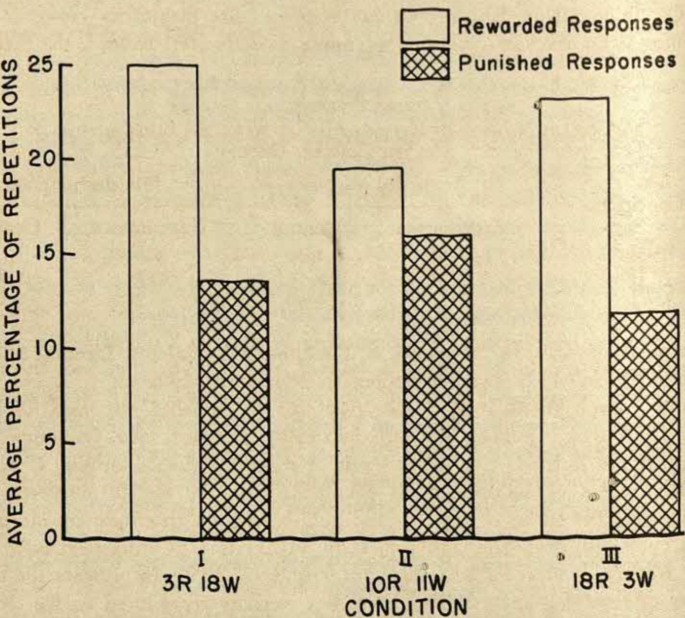


FIG. 1. AVERAGE PERCENTAGES OF REWARDED AND PUNISHED RESPONSES REPEATED ($R =$ right; $W =$ wrong)

of this difference reaches a satisfactory level of significance (for Condition III in which rewarded responses were highly crowded). The direction of the results is, however, the same in all cases. When all conditions are combined, rewarded responses are repeated more frequently than punished ones at a high level of significance ($t = 3.12$, $P < 0.01$). Analysis of variance shows that the differential effectiveness of rewards and punishments does not vary significantly from condition to condition ($F = 1.31$, *df* 0.2 and 57, $P > 0.05$).

(4) Relation between isolation and repetition. The most conservative conclusion from the analysis of the repetitions is that the effects of rewards and punishments

are independent of the factor of perceptual isolation. None of the comparisons yields a significant difference among conditions. In spite of the lack of statistical significance there is, however, a consistent tendency for both rewards and punishments to be less effective in Condition II, *i.e.* when both kinds of reinforcement occur with about equal frequency. Assuming on the basis of the consistency of the results that this finding represents a real trend, it appears that the frequent alternation of rewards and punishments serves to reduce the effectiveness of the rewards. Such alternation probably results in a considerable amount of intraserial interference, *i.e.* a failure to discriminate between rewarded and punished pairs. To speak of intraserial interference is not just to use another phrase for 'perceptual crowding.' If the effectiveness of reward did, indeed, depend on perceptual isolation, Condition III in which rewarded responses are maximally crowded should yield the lowest level of repetition of rewarded responses. Instead, the percentage

TABLE II

AVERAGE NUMBER OF RESPONSES RECALLED BY THE EXPERIMENTAL AND THE CONTROL GROUPS

Condition	Rewarded responses		Punished responses	
	Experimental	Control	Experimental	Control
I	1.05	1.00	3.90	5.05
II	2.35	2.75	2.25	2.80
III	5.10	5.50	.70	.65

of repetition of rewarded responses in Condition III is almost identical with that in Condition I in which rewarded responses are maximally isolated.

Similar considerations apply to the repetition of punished responses. Both Condition I, in which punishments are maximally crowded, and Condition III in which they are isolated, give lower levels of repetition than Condition II in which the effectiveness of punishment is counteracted by intraserial interference. We conclude that the effects of rewards and punishments on repetition are independent of perceptual isolation although subject to modification by intraserial interference.

(5) Recall of rewarded and punished responses. Table II presents the average number of rewarded and punished responses correctly reproduced on the recall-test by all experimental and control groups. For purposes of comparison the same results are presented in percentage form in Fig. 2.

The recall performances of the experimental and control groups are quite similar. None of the experimental groups differs significantly from its control group in the recall of either rewarded or punished responses.

For both experimental and control groups, there are only small and insignificant differences in recall as a function of isolation. There is some indication that isolated rewarded responses (Condition I) are recalled better than highly crowded ones (Condition III). The fact that Condition II yields poorer retention than either of the other conditions makes it difficult, however, to ascribe even this trend to the factor of isolation. In the case of the recall of punished responses, no consistent trend emerges.

In agreement with the results of our previous study, rewarded responses are

recalled better than punished ones by all groups. For all experimental and control groups combined, the difference is highly significant ($t = 3.57$, $P < 0.01$). There is no significant variation among groups in the degree to which rewarded responses were favored over punished ones in recall.

(6) Identification of reinforcements. On the recall-test, *S* indicated for each

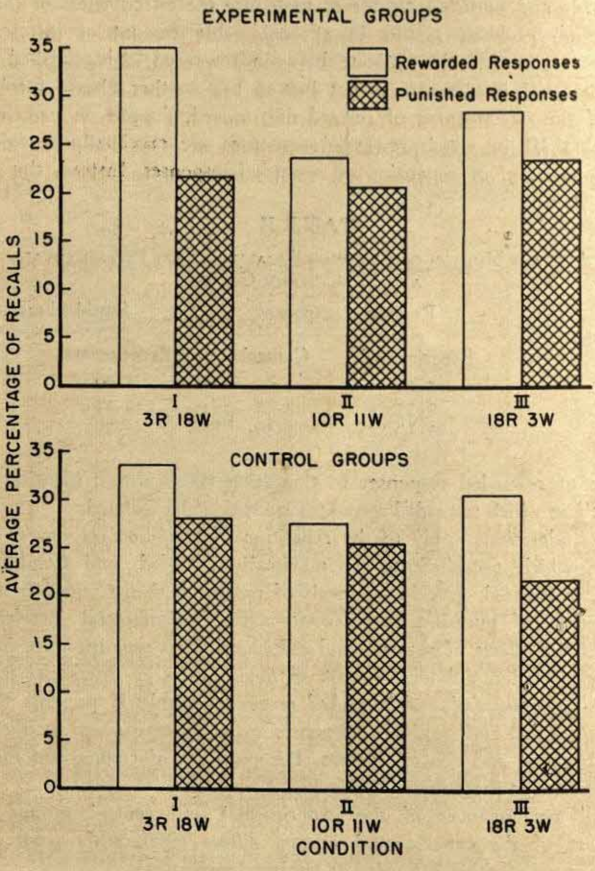


FIG. 2. AVERAGE PERCENTAGES OF REWARDED AND PUNISHED RESPONSES RECALLED ($R =$ right; $W =$ wrong)

response whether it had been rewarded or punished during the training series. To make the results of different conditions comparable, the opportunities for correct chance guesses offered by the different series must be taken into account. Correction for chance successes is made by subtracting the percentage of incorrect identifications from the percentage of correct identifications. The resulting accuracy scores for all experimental and control groups are presented in Table III. For both ex-

perimental and control *Ss*, accuracy of identification is considerably higher in Condition I, in which rewards were 'isolated,' than in Conditions II and III. None of the experimental groups is significantly different from its control group. Analysis of variance of the identification scores of all groups shows the difference among conditions to be highly significant ($F = 9.40$, $df = 2$ and 114, $P < 0.01$). Further tests of significance reveal that the critical break occurs between Condition I and the other two conditions.[11]

Isolation (or reduction in intraserial interference) improves the accuracy of identification only when the number of rewards is small relative to the number of punishments. Apparently it is the rewards, rather than the punishments, to which the *Ss* respond differentially during training and which provide them with landmarks for identifying the positions of reinforcements. It should be noted, however, that *Ss* tend to underestimate the total number of rewards which they have

TABLE III

ACCURACY OF IDENTIFICATION OF REINFORCEMENTS BY THE DIFFERENT EXPERIMENTAL AND CONTROL GROUPS

(Entries are percentages correct minus wrong.)

Condition	Percentage accuracy	
	Experimental	Control
I	19.17	22.50
II	7.14	−9.31
III	6.11	1.39

received, and the larger the number of rewards the greater is the degree of this underestimation. This tendency serves to increase the percentages of error, particularly in Condition III.

(7) Gradient of identification of rewards. The fact that rewards serve as anchor points in the discrimination of the positions of reinforcement is supported further by analysis of the incorrect identifications. In Condition I reinforcements were so spaced that each reward was preceded and followed by three punished responses; in Condition III the punishments were spaced in a like manner. In these two cases it was possible to determine, therefore, by how many steps an item recalled as rewarded (Condition I) or recalled as punished (Condition III) preceded or followed the one actually reinforced in the training series. When the frequency of identifications is plotted against distance from the actually reinforced response, the results shown in Fig. 3 are obtained. (Since the results for the experimental and control groups were quite similar, they were combined for purposes of this presentation.)

The frequency of identification of rewards describes a gradient with respect to the actually rewarded position, in agreement with the findings in our previous experiment.[12] In the case of punished responses (Condition III) errors of identifica-

[11] Determining the critical break between conditions, we followed the procedure described by J. W. Tukey, Comparing individual means in the analysis of variance, *Biometrics*, 5, 1949, 99-114.

[12] Postman and Adams, *op. cit.*, 612-631.

tion have no systematic relationship to distance from the actual reinforcement.[13]

(8) Isolation and recall. The *Ss'* recall of their own responses is independent of the crowding or isolation of the *reinforcements*. The correct recall of the exact positions of the rewards and punishments, on the other hand, does depend on the relative frequency with which the two kinds of reinforcements appear in the series:

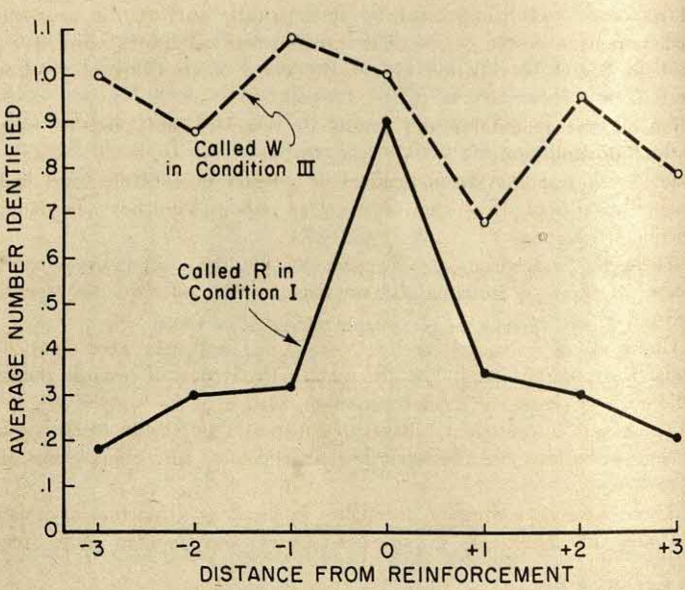


FIG. 3. AVERAGE NUMBER OF RESPONSES CALLED 'RIGHT' AND 'WRONG' ON THE RECALL-TEST AS A FUNCTION OF THE DISTANCE FROM THE POSITION OF PREINFORCEMENT

(Combined data for experimental and control groups.)

the less frequent the *rewards* the more accurate is the identification of reinforcements.

These findings are in agreement with the general hypothesis that the characteristics of the series influence performance only to the extent that they control the amount of intraserial interference. The series of items which the *S* has to *recall* are equally homogeneous under all conditions, *i.e.* they all consist of 21 *numbers*. In the case of the *identification of reinforcements,* on the other hand, opportunity for intraserial interference increases with the number of *rewards*. It is not surprising, therefore, that the distribution of reinforcements has different effects on the two measures of retention.

[13] Evidence for a 'spread' or effect around the rewarded positions was found in Condition I. There was no gradient of repetition around the punished positions in Condition III. Only very slight evidence for spread of repetition around an isolated wrong response was reported by Zirkle, *op. cit.*, 304.

Summary and Conclusions

This experiment was designed to evaluate the relative importance of reinforcement and perceptual isolation as determinants of response repetition in a Thorndikian situation. A series of 21 words was presented to each of which *Ss* responded with a number from 1 to 10. Word-number associations were differentially reinforced by announcements of 'right' and 'wrong.' The degree of isolation of rewarded responses was manipulated by systematically varying the proportion of rewarded responses in the series. The experimental conditions covered a range from a high degree of isolation to a high degree of crowding of rewarded responses.

Analysis of the repetitions and recalls of rewarded and punished responses yielded the following results:

(*1*) Rewarded responses were repeated at a better than chance level under all experimental conditions. Repetition of rewarded responses did not vary significantly as a function of isolation.

(*2*) There was a tendency for repetition of punished responses to exceed the chance-level. Degree of isolation did not significantly influence the repetition of punished responses.

(*3*) Under all experimental conditions, rewarded responses were repeated more frequently than punished ones. The differential effectiveness of rewards and punishments did not vary significantly as a function of isolation.

(*4*) There was a consistent, although statistically unreliable, tendency for reinforcements to be least effective when rewards alternated with punishments in about equal proportions.

(*5*) There were no significant differences in recall as a function of isolation.

(*6*) Under all conditions, rewarded responses were recalled more frequently than punished ones.

(*7*) Accuracy of recall for the position of reinforcements in the series varies significantly with the proportion of rewards. The smaller the proportion of rewards the greater is the accuracy of identification.

(*8*) The frequency of identification of rewards describes a symmetrical gradient with respect to the actually rewarded position. For punished responses, errors of identification have no systematic relationship to distance from the actual reinforcement.

We conclude that the effectiveness of reinforcements is independent of the isolation of rewards and punishments in the series. Thorndike's interpretation of the effects of 'right' and 'wrong' cannot be adequately challenged by reference to the effects of perceptual isolation.

Our results highlight the need for precision in applying concepts such as 'isolation' to the analysis of learning data. In previous attempts to explain Thorndike's findings in terms of a perceptual analogue, no clear distinction was made between the isolation of responses, the isolation of reinforcements, and the isolation of stimulus-response associations. Our findings suggest that it is only in the recall of the positions of reinforcements that isolation of rewards has an effect; it does not significantly influence either the repetition or the recall of responses. Whatever effects are found can be adequately accounted for in terms of intraserial interference.

DETERMINANTS OF CHOICE-DISTRIBUTION IN TWO-CHOICE SITUATIONS

By JACQUELINE JARRETT GOODNOW, Harvard University

When one is faced with a series of choices between two alternatives, each of which involves probable rather than certain outcomes, what are the conditions that determine whether one will split his choices between the two alternatives or will concentrate on one of them? Take, for example, a situation in which one has to predict on each trial which of two possible events will occur when the series of trials has been randomized and so arranged that the probability of event A occurring is 0.7 and of event B, 0.3. Choices in such situations have consistently been found to be split between these alternatives.[1] In terms of the example, one learns to predict event A on roughly 70% of the trials and event B on roughly 30% of them. For the sake of convenience, this form of choice-distribution shall be called here 'agreement between choices and event-probabilities.' Despite the consistency of such results, the conviction remains that a 100:0 distribution of one's choices—always predicting the event more likely to occur—is a highly reasonable form of behavior and that any theory dealing with choice-situations must either explain why it does not occur or else state the experimental conditions under which it would occur.[2]

The present experiment tests a condition suggested by game-theory; namely, the relation between a single win or loss and the final outcome at which the player aims. A matched pair of two-choice situations (one

---

* Accepted for publication April 9, 1954. This paper was adapted from a doctoral dissertation submitted to Harvard University and directed by Dr. J. S. Bruner. The study was conducted in 1950 and supported by the Mathematical Models Project, Laboratory of Social Relations. The author's thanks are due Dr. Frederick Mosteller and Dr. R. R. Bush for advice on analysis and for the apparatus used in the gambling experiment.

[1] See W. O. Jenkins and J. C. Stanley, Jr., Partial reinforcement: A review and a critique, *Psychol. Bull.,* 41, 1951, 291-297; R. R. Bush and Frederick Mosteller, *Stochastic Models for Learning,* (in press) Chap. 13.

[2] The problem has been discussed by D. A. Grant, H. W. Hake, and J. P. Hornseth, Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement, *J. Exper. Psychol.,* 42, 1951, 1-5; H. W. Hake and Ray Hyman, Perception of the statistical structure of a random series of binary symbols, *ibid.,* 45, 1953, 64-74.

problem-solving and the other gambling) was designed which differed only in the relation between a single win or loss and the desired final outcome. The hypothesis set up was that 'agreement-type' distribution of choices would be found in 'problem-solving,' and that the 100:0 type or something close to it would occur in gambling. The steps leading to this hypothesis are as follows.

The selection of relation between a single win or loss and the desired final outcome as an experimental variable stemmed from Von Neumann and Morgenstern's discussion of factors important in the development of strategy; especially from their statement that the effect of a loss, risked or sustained, is relative to the total outcome.[8] One can regard a partial reinforcement as a form of 'risk' or 'uncertainty' and expect that at least some of the variables found important in game-theory should be relevant to a two-choice problem in learning. From such a position, a problem-solving, two-choice task can be analyzed as follows:

(*a*) The desired final outcome is 100% success, *i.e.* eventually solving the problem so that one can predict correctly on every trial. Anything less than this is not really success.

(*b*) The final outcome is not determined by a count of wins and losses. On the contrary, finding the solution at the end makes unimportant all previous errors. In this sense, one can afford errors while searching for the correct principle.

(*c*) Under such conditions, 100:0 choice-distributions should not occur. In a 70:30 situation, for example, the 100:0 choice-distribution allows one to be correct only 70% of the time and this cannot be considered as solving a problem. Logically, the only kind of choice-distribution which allows one to be correct on every trial is one which matches the probabilities of the events, *i.e.* a 70:30 distribution in a 70:30 situation, a 90:10 distribution in a 90:10 situation, and so on.

The best check on the usefulness of this type of analysis is, of course, that of setting up a task where the relation between a single win or loss and the final outcome is different from that involved in problem-solving and is conducive to a 100:0 type of choice-distribution. A 'gambling task,' regarded as meeting these requirements, is the following.

(*a*) The desired final outcome is not necessarily 100% success. One can be successful without having to know how to win on every trial.

(*b*) The final outcome is determined by a count of single wins and losses. The task can be so set that a win at the end cannot wipe out previous losses. All that this involves is the restriction of having the same amount of money involved in each bet and in each payoff. In this sense, there are limits placed on the extent to which one can afford losses.

[8] John von Neumann and Oscar Morgenstern, *Theory of Games and Economic Behavior*, 1944, 163.

(*c*) Under Conditions (*a*) *and* (*b*), choice-distributions should tend to be of the 100:0 type. In a 70:30 case, for example, an 100:0 choice-distribution will allow one to win on 70% of one's bets and this can easily be regarded as 'success.' Also, the probable losses incurred by choosing the 30% alternative cannot be well afforded, since every single loss counts in the final score. It is not to be expected, however, that the group mean choice-distribution will completely reach 100:0. Choosing the same alternative all the time is rather boring, especially when the stakes are small. Furthermore, there will always be some players who aim for a 'system' and try to win on every trial, no matter what instructions are given. Still others will be attracted away from 100:0 by the subjective value placed upon the act of gambling itself,[4] and will find the amount of money involved a poor deterrent to making a bet on the 'riskier' alternative.[5]

Method and Procedure

Apparatus: (1) Problem-solving. This task has been described in detail in a previous paper.[6] Briefly, it is an insoluble task with a variety of geometrical designs as material. On each trial *S* is presented with three cards, each containing a two-figure geometrical design. The first card contains a key-design. On each of the other two cards is a variation of the key-design. *S*'s task is to couple the key-card with one of the variation cards and to find the principle by which the coupling should be made.

To form the variable cards, the key-design is altered either by adding or subtracting a line. *S* is, therefore, choosing on every trial between the two types of variation. It is these two types that are the real alternatives—the equivalents of left and right turns in the T-maze. The array of specific geometrical forms which the key-designs and their variations may take is the problem-solving disguise. With 10 different key-cards and 4 variations for each of these, the disguise is both elaborate and convincing enough to make *S* believe that he is working on a soluble but complex problem. *S* is led to believe that his choice should be determined by some hypothesis about the nature of the specific designs, and that the success or failure of his choice is dependent upon his ability to find the correct hypothesis. Actually, the success or failure of his choice is dependent upon whether or not it agrees with *E*'s predetermined schedule of 'correct' answers. In a 70:30 case, for instance, the schedule is such that on a random 70% of the trials, one type of variation is set up as the correct choice to make, while the other type of variation is the correct choice to make on the remaining 30% of the trials.

(2) *Gambling.* The gambling task involves playing a slot machine with chips worth one cent each (see Fig. 1). *S* is staked by *E* to 200 chips. He inserts one of these chips into the machine and presses one of two keys. If the bet is successful, a chip drops into the payoff box with a clatter of noise. The payoff box has a glass face and *S* can see the heap of chips he has won. He is not allowed to take

[4] This point has been especially made by M. G. Preston and Philip Barratta, An experimental study of the auction-value of an uncertain outcome, this Journal, 61, 1948, 183-193.

[5] Frederick Mosteller and Philip Nogee, An experimental measurement of utility, *J. Polit. Econ.* 59, 1951, 371-404.

[6] Jacqueline Goodnow and Leo Postman, Learning in a two-choice probability situation with a problem-solving setting, *J. Exper. Psychol.*, 49, 1955, 16-22.

the chips out of this box until the end of the experiment. Whatever the outcome of the bet, the machine becomes inoperative for several seconds between each trial, and *S* must wait until the appearance of two signal lights and a loud buzz which indicate that the apparatus is ready to take the next bet. The apparatus is so wired that it is useless for *S* to insert a chip before the machine is ready; to press a key before inserting a chip; or to press two keys on any one trial.

Lights on the face of the machine flash on successively as it comes into operation (upper outer lights in Fig. 1), as *S* inserts a chip (upper center light), and

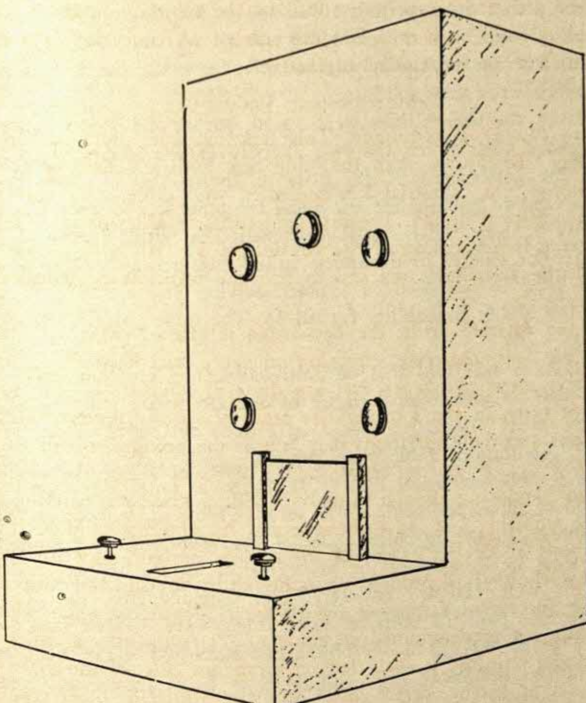


FIG. 1. OUTLINE DRAWING OF THE APPARATUS USED IN THE EXPERIMENTS IN GAMBLING

as he presses a key (left or right lower light). These lights are paralleled by lights on the control machine operated by *E* in an adjoining room. By watching the lights on the control machine, *E* knows when *S* has bet and which key he pressed. The control machine also contains a master switch to turn the apparatus on or off, and a key which, when pressed, causes the machine to eject a chip into the payoff box. A one-way mirror allows *E* to observe *S* from the control room.

Some of the features of the machine, *e.g.* the number of lights, have as their primary purpose that of helping *S* to believe that the operation is automatic. To

strengthen this belief, *S* was given a little information about relays, about the machine 'operating on a punched-tape system,' and about the use of a one-way mirror for making observations and keeping time-scores without affecting behavior. He was also asked to delay making the first bet until *E* tapped on the one-way mirror, that *E* would 'have time to get settled at the observation table.' Judging from the statements of the *Ss* at the conclusion of the experiments, most of them accepted these explanations.

Procedure. Procedure and instructions for the two tasks were designed to parallel one another as far as possible. The only real difference was that in problem-solving *S* was informed that the experiment was one in learning and that he was to look for a principle, whereas in the gambling task he was told that he was playing for money in an experiment on gambling and that he would be paid for the difference between the number of wins and losses.

In both tasks, the probabilities were 50:50, 70:30, and 90:10. Probabilities refer to the frequency with which the two alternatives are scheduled to pay off. Case 70:30 with the gambling task, for instance, means that one key is scheduled to pay off on a random 70% of the trials and the other key is scheduled to pay off on the remaining 30% of the trials. Randomization is within blocks of 10 trials. On the basis of pilot studies to determine initial preferences for one alternative, the 'subtracting' type of variation in problem-solving and the left key in gambling were given the higher probability for all *Ss*. The same payoff schedules were used for both tasks.

In both tasks *S* was told that one alternative and only one would pay off on every trial—that an unsuccessful choice of one alternative automatically meant that choice of the other alternative would have been successful. Every *S* was allowed 120 trials, divided into 12 blocks of 10 trials each. Inter-trial and inter-block intervals were respectively 3 and 10 sec. *S*'s choices in both tasks were recorded by *E*.

At the end of the experiment, *S* was asked: "How did you decide which alternative to choose?" In the gambling task, if the question was appropriate, he was also asked what he thought of the procedure of always betting on one key.

Subjects. Ss were Harvard undergraduates obtained from the Student Employment Bureau and paid 75 cents an hour. They were divided into 6 groups, and every *S* in a group was given one of the two tasks and one of the three probabilities: 50:50, 70:30, and 90:10.

RESULTS

Results are presented in terms of mean proportion of choices of one alternative: 'subtracting' in the problem-solving, and 'left key' in the gambling groups—these alternatives being given the higher probability outside the 50:50 case. Results are presented in Table I.

Table I indicates that there is no difference between gambling and problem-solving groups for case 50:50, a sizable difference between the two groups for case 70:30 and a possible difference between the two groups for case 90:10. Statistical analysis supports this inspection. The Mann-Whitney test was used, the score for each *S* being the number of

times *S* chose 'subtracting' or 'left' over the last 50 trials of the 120 trial learning series.[7] This technique allows us to reject at the 95-% confidence level the null hypothesis that the gambling and problem-solving groups for case 70:30 are identical. It does not allow us to reject the hypothesis that gambling and problem-solving groups are identical for cases 50:50 and 90:10.

Case 90:10, however, has some interesting features which need to be noted. In the problem-solving group, the number of choices of the 90% side over the last 50 trials by the 10 *Ss* is as follows: 39-43-45-45-46-47-47-48-50-50. In contrast, in the gambling group, the number of choices of the 90% side over the last 50 trials by the 14 *Ss* is as follows: 41-42-42-47-48-49-50-50-50-50-50-50-50-50. It is clear that with the limited possibility for variation in scores, and with the within-group differences obtained in samples of the size used here (10 and 14), it is difficult to obtain a statistically significant difference for a 90:10 case. The trend towards a difference certainly exists but larger samples would be needed to give the difference statistical significance. In general, the conclusion can be made that, outside the 50:50 case, the use of a gambling task gives rise to a higher concentration of choices of one alternative—a closer approximation to a 100:0 choice-distribution—than does a problem-solving task.

TABLE I

MEAN PROPORTIONAL CHOICES BY GROUP AND BY BLOCKS OF 20 TRIALS OF 'SUBTRACTING' IN PROBLEM-SOLVING AND 'LEFT KEY' IN GAMBLING

Trials	Group 50:50		70:30		90:100	
	Problem-solving	Gambling	Problem-solving	Gambling	Problem-solving	Gambling
1–20	.470	.430	.485	.489	.675	.765
21–40	.475	.505	.510	.664	.910	.878
41–60	.455	.550	.585	.721	.925	.950
61–80	.500	.495	.625	.722	.920	.954
81–100	.515	.465	.590	.782	.885	.965
101–120	.505	.515	.665	.815	.955	.964
N	10	10	10	14*	10	14*

* For these two groups *N* was increased to 14 because the data were required for another purpose (an analysis of choice-sequences) and the incidence of 100:0 choice-distributions was cutting down on the amount of data available for such analysis.

[7] H. B. Mann and D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, *Annals. Math. Stat.* 18, 1947, 50-60.

Discussion

In the original formulation of this experiment, the relation between a single win or loss and the desired final outcome was postulated as an important condition in the occurrence of agreement or 100:0 type of choice-distribution. On the basis of the detailed experimental results (*Ss*' statements and observations of their behavior), this condition is seen as still of major significance but as an over-simplification of the problem. 'Desired final outcome,' for example, is still an important goal but needs to be supplemented by a statement of sub-goals such as those of proving one's skill by correctly predicting the infrequent event, and obtaining direct information about the occurrence of events. 'Single wins' need to be considered not only in terms of the relation to the final outcome but also in terms of their relation to these sub-goals. 'The relation between single losses and the final outcome' is more happily regarded as a particular form of a more general condition, *i.e.* the extent to which *S* can afford and can overlook losses.

A fuller description of these and a few subsidiary conditions and the way they affect the choice-distribution is summarized as follows.[8]

(1) Conditions leading to 100:0 choice-distributions: (*a*) The acceptance of less than 100% success as a good final outcome. This implies that even in the final stage of the game or learning series a limited amount of loss or error is acceptable.

(*b*) The lack of belief in a system which will allow 100% success or will at least make advisable a departure from a safe and sure method with the risk such

---

[8] These are certainly not the only conditions affecting the choice-distribution. One must draw from learning theory to account for the way in which *S* learns or acquires information about the event-probabilities. The conditions discussed here concern *S*'s utilization of information rather than its acquisition. We know that the *Ss* in both problem-solving and gambling groups give, when asked at the end of the experiment, estimates of event-probabilities so similar that the difference between the two groups does not lie in their knowledge of probabilities. The difference seems to lie in the way this knowledge combines with such conditions as an aim of eventual 100% success or ability to afford loss and results in a particular choice or decision.

The conditions stated above can also be drawn together in terms of a useful general condition provided by the Bush-Mosteller model for learning. Bush and Mosteller distinguish between two-choice situations in terms of whether or not a rewarded A_1 response has nearly the same effect on future behavior as a non-rewarded A_2 response and vice versa. The conditions stated in this paper can be regarded as psychological conditions which affect the way in which a rewarded or non-rewarded A_1 response influences later responses. It can be argued, for instance, that a non-reward on the 70% key will not increase the probability of choice of the 30% key if *S* has already accepted the possibility of loss on the 70% side or if *S* considers that occasional payoffs on the 30% side tell him nothing about how to win on the 30% side on a later trial.

departure involves of increasing one's losses beyond the amount anticipated and allowed for.

(*c*) The ability to overlook losses. In the gambling task, for instance, getting no return for one's chip seems to be a negative rather than a positive reminder of loss and, as long as such a reminder does not occur too often, *S* tends not to become disturbed about the possible wins he is allowing to pass by.

(*d*) The ability to set up a substitute activity which will relieve the boredom of always doing the same thing and will enable one to resist the appeal of taking a risk. This is best illustrated by two *S*s in the 90:10 gambling group. 'To make things interesting' while always betting on the left key, they set up new problems such as determining whether or not the amount of pressure on the hand used to press the left key made any difference to the outcome of the bet. Both *S*s stated that they knew these were 'fake' problems but that some such game was necessary to keep them from switching to the 'more interesting' right key.

(2) Conditions leading to other choice-distributions: (*a*) An interest in eventual 100% success. This implies the hope that loss or error incurred during the game or the learning series will be made insignificant by perfect performance in the final stages.

(*b*) An interest in a level of success which, while not reaching 100%, will possibly be higher than the minimum allowed by 100:0 behavior. This is strengthened by the attitude that odds of 70:30 and 90:10 are so 'good' that at the worst one will not fall far below this minimum and losses are again affordable.

(*c*) An interest in proving one's skill by correctly predicting the infrequent event. This was particularly noticeable in the gambling task where many *Ss* preferred to win on a lesser total number of bets if they could have the satisfaction of winning on the right key with its lower probability of paying off, even though the amount of money to be won from either key on a bet is the same. The source of the satisfaction seems to lie in the idea that one has been smart enough or lucky enough to win in the face of the odds. In the words of one *S:* "It's easy to win on left; the real skill comes in winning on right." It should be noted that, like the goal of being correct on every trial, this goal has the effect of bringing the choice-distribution into line with the event-probabilities. One could win on the right key by betting on it very often but—apart from the fact that this involves a reduction in the total number of wins—this would be no more a test of skill than deer-hunting with a machine gun. The real skill lies in betting on the right key just on those occasions when it is likely to pay off.

(*d*) A belief in the existence of a pattern or system which will allow one to achieve eventual 100% success or predict the infrequent event.

(*e*) An interest in obtaining direct information about an event. The importance of this interest was quite unexpected, since the experimental situation was such that choice of one alternative automatically yielded information as to what would have happened if the other alternative had been chosen. *S*s could check hypotheses about the 'riskier' alternative while staying with the 'safer' alternative, but most *S*s find this an unsatisfactory procedure. They state that the procedure is "too indirect," that they "would miss out on actually having been correct," and that "the only real way to check a hypothesis about 'subtracting' is to choose it and see what happens." It would seem that the effect of this interest depends on *S*'s ability or

readiness to utilize direct information and on the hope that there may be a way of doing better than he is doing now. The phenomenon is rather similar to *S*'s reluctance or inability to use negative instances in studies on concept-formation.[9]

(*f*) Limits on the ability to overlook losses. In the problem-solving task, for instance, *S* can afford loss while searching for direct information or a system. He cannot, however, overlook losses or errors, partly because of his general set to find eventually a way to eliminate errors and partly because his concern with the nature of the geometrical design which was correct constitutes a fairly positive reminder of error.

These conditions may also be applied to 'risks' or 'uncertainties' other than the kind investigated in the present study. In gambling, for instance, *S*s are often offered a choice between bet or no bet, or between the acceptance of different sets of odds for diffrent amounts of money. For these situations, the presence of goals other than those of 'maximizing expected wins' and 'minimizing expected losses' (namely, such goals as aiming at a system and eventual 100-% success or proving one's skill by predicting an event in the face of the odds) may help to explain—in terms of utility, psychological probability, and probability-preferences—why the *S*s show differential responses to risk and do not seem to follow the model suggested by game-theory.[10]

How do the conditions discussed apply to two-choice situations other than the kind investigated in the present experiment? It is suggested that the experimental tasks used in studies where 'agreement-type' choice-distributions have been found can easily be regarded as essentially tasks in problem-solving in which the *S*s are aiming at being correct on every trial, or at proving their skill by correctly predicting the infrequent event. This is certainly true of tasks of the Humphreys'-type, from which most of the evidence has come. Here *S*s are required to predict whether a second light will appear following the appearance of a first light and they are either explicitly instructed or set "to do their best to answer correctly."[11]

---

[9] See C. I. Hovland and Walter Weiss, Transmission of information concerning concepts through positive and negative instances. *J. Exper. Psychol.*, 45, 1953, 175-183.

[10] The discrepancy between experimental data and the game-theory model has recently been pointed up by Ward Edwards, Probability-preferences in gambling, this JOURNAL, 66, 1953, 349-364. Discussions of behavior in economic risk and uncertainty situations are given by George Katona, Rational behavior and economic behavior, *Psychol. Rev.*, 60, 1953, 307-318; L. J. Savage, The theory of statistical decision, *J. Amer. Stat. Assoc.*, 46, 1951, 55-67; and C. H. Coombs and D. C. Beardsley, Decision making under uncertainty, in R. M. Thrall, C. H. Coombs, and R. L. Davis, (eds.), *Decision Processes*, (in press).

[11] These are the instructions reported by L. G. Humphreys, Acquisition and

The same conditions can also be applied to two-choice situations where the probabilities do not add up to unity or where there is not the requirement that one and only one event can occur on each trial. In such two-choice situations, approximations to 100:0 choice-distributions have been found more frequently although not consistently.[12] It is suggested that this result is due to *S*'s inability to regard the situation as one amenable to a problem-solving approach or the use of skill. Where there is not the requirement that one and only one event can occur on each trial, *S* is essentially faced on every trial with four possibilities, *e.g.* key 1, key 2, both keys or neither key could turn on a light. As long as *S* is restricted to choosing only one key on each trial, whatever choice he makes can eliminate only one of the four possibilities. *S* can never acquire sufficient information to make a search for a system or pattern worthwhile and choice of the risky alternative becomes very 'risky' indeed. When in addition to such incomplete information, *S* is faced with probabilities which do not add up to 100:0, the inevitability of a certain amount of loss becomes even more obvious. Also, losses or errors lose much of their sharpness when they are so inevitable and when they tell *S* so little about what else he could have done. All in all, the temptation is strong for *S* to regard the situation as an 'unbeatable chance' one in which the best one can do is to keep the amount of loss to a safe minimum by 100:0 behavior. This is admittedly an overstatement of what happens in such two-choice situations, but it is suggested as one possibly fruitful way of drawing together data from different types of two-choice situations.

SUMMARY

This experiment compared performance in a matched pair of two-choice situations—a problem-solving task and a gambling task. Each of the 6 groups of *S*s was given one of the two tasks and one of three event-probabilities: 50:50, 70:30, or 90:10. In the 50:50 groups for

---

extinction of verbal expectations in a situation analogous to conditioning, *J. Exper. Psychol.*, 25, 1939, 298. They were presumably also used in other experimental studies replicating Humphreys' task. Similar instructions are given by M. E. Jarvik, Probability learning and a negative recency effect in the serial anticipation of alternative symbols, *J. Exper. Psychol.*, 41, 1951, 291-297. In Jarvik's study, *S*s were required to predict whether *E* would say 'plus' or 'check.'

[12] The most extensive study with this kind of two-choice situation has been by E. D. Neimark, Effects of type of non-reinforcement and number of alternative responses in two verbal conditioning situations, Ph.D. thesis, Indiana University, 1953. This study and several other unpublished studies are reviewed by Bush and Mosteller, *ibid.*, chap. 13.

both tasks, *Ss* distributed choices between alternatives in a 50:50 fashion. In the 70:30 and 90:10 groups, the gambling *Ss* distributed choice in more of a 100:0 fashion than did the problem-solving *Ss*. This difference is explained predominantly on the basis of *Ss*' goals and interests (whether the task sets a final goal of 100% success or less than 100% success is a major factor), and the extent to which the task enables *S* to afford and to overlook losses.

A SCALE FOR MEASURING SUPRA-THRESHOLD OLFACTORY INTENSITY

By LAWRENCE KRUGER, ALVIN N. FELDZAMEN, and WALTER R. MILES,
Yale University

The problem of defining and classifying olfactory intensity is complicated by the ambiguous position of odorants on the wide olfactory spectrum. It is well-known, for example, that changing the concentration of an odorant, and thereby its intensity, may change its quality as well, and that odorants of different quality may serve to mask each other. At the present time there is no general agreement as to the physico-chemical properties of odorants which could serve as a basis for a satisfactory odor classification.

It would be possible, if primary qualities covering the odor spectrum existed or if the relevant physico-chemical variables could be determined, to attempt the construction of a scale of intensities for the different qualities. Lacking this knowledge, one can choose an arbitrary odor as a standard and use that odor for the construction of a scale of intensities, as in vision and audition. To construct such a scale, matching procedures against an arbitrary stimulus are required. An accurate scotopic visibility curve, for example, was finally obtained only by matching lights of various wave-lengths against a 'white' light of fixed low intensity, and an intensity scale was constructed by varying the intensity of the 'white' standard light.[1] In the study of the pitch-intensity problem in audition it was found necessary to match, for intensity, tones of different pitch against an arbitrary tone of 1000 ~ at several intensive levels.[2] Although pitch changes as the loudness varies, this method led to a consistent and useful intensive scale.

* Accepted for publication February 15, 1954. This study was supported by grants-in-aid from Arthur D. Little, Inc., Cambridge, Massachusetts, and The Research Corporation, New York City. The authors wish to express their thanks to Dr. L. H. Beck, Department of Psychology, Yale University, for his advice and encouragement, and to Dr. P. A. Lyons and Dr. J. M. Sturtevant of the Sterling Chemistry Laboratory, Yale University, for providing space and guidance for the measurement of vapor pressures.

[1] Selig Hecht and R. E. Williams, The visibility of non-chromatic radiation and the absorption spectrum of visual purple, *J. Gen. Physiol.*, 5, 1922, 1-34.

[2] Harvey Fletcher and W. A. Munson, Loudness, its definition, measurement, and calculation, *J. Acoust. Soc. Amer.*, 5, 1933, 52-108.

An early attempt to assign numerical values to supra-threshold intensities was provided by the development of the Zwaardemaker olfactometer.[3] With this apparatus units of intensity are expressed in terms of an odorant's area of exposure. The method suffers from the disadvantage that its data are not comparable because 10 olfacties of numerous substances would vary considerably in their respective olfactory intensities. In an attempt to evaluate gas warning agents, Allison and Katz[4] and Fieldner and his co-workers[5] developed a scale which consisted of matching an unknown odorant against a subjective description ranging from zero, 'no odor,' to 5, 'very strong.' Crocker and Henderson, interested in the quality of odors as well as their intensity, introduced a system of classification based on eight intensive levels in each of four primary qualities ('fragrant,' 'acid,' 'burnt,' 'caprylic'), which, despite its potentialities for perfume chemistry, suffers many theoretical disadvantages.[6]

Most quantitative data gathered in the study of smell have been obtained by threshold measurements. Many experimental difficulties, inherent in these measurements, have received attention from workers in this field.[7] The increased effect at threshold levels of the stimulus-contaminants, which are almost always present, and the variations in the results encountered with different methods of introducing the odorants in the nose, are serious experimental problems. While these difficulties may be minimized by further research, the question regarding the usefulness of threshold data still remains. The failure of many odorants, detectable at low concentrations, to be intense at high concentrations, indicates that there is no simple relation between threshold and supra-threshold intensities. Threshold values cannot be successfully used for predicting olfactory intensity. For example: musk is detectable at a remarkably low concentration, but it never, even at very strong concentrations, becomes as powerful as odors whose thresholds are many times higher. It would seem, then, that a systematic study of supra-threshold intensities would furnish more significant information than further studies of the threshold.

It is proposed that olfactory intensity may be treated in a manner analogous to that used in visual and auditory studies by the comparison of odorants with a 'standard' odorant, the analogue of the fixed 'white' light and the fixed tone. We assume that qualitative changes may be ignored and odorants matched for intensity alone against a standard. This problem also exists in vision, where the observer (*O*) must learn

---

[3] Hendrik Zwaardemaker, *L'Odorat,* 1925, 305.

[4] V. C. Allison and S. H. Katz, An investigation of stenches and odors for industrial purposes, *J. Ind. Eng. Chem.,* 11, 1919, 336-338.

[5] A. C. Fieldner, R. R. Sayers, W. P. Yant, S. H. Katz, J. B. Shohen, and R. D. Leitch, Warning agents for fuel gases, *U. S. Dept. Commerce, Bureau Mines Monog.* No. 4, 1931, 168.

[6] E. C. Crocker and F. L. Henderson, Analysis and classification of odors, *Amer. Perfumer,* 22, 1927, 3-10; E. G. Boring, A new system for the classification of odors, this JOURNAL, 40, 1928, 345-349.

[7] B. M. Wenzel, Techniques in olfactometry: A critical review of the last one hundred years, *Psychol. Bull.,* 45, 1948, 231-247; R. W. Moncrieff, *The Chemical Senses,* 2nd ed., 1951, 97-112.

to ignore the color of his 'white' light standard. A shift in spectral matching accompanies a shift in the brightness of the 'standard' lamp. These variables can only be determined by specifying the color temperature and the brightness of the lamp used for comparison. In olfaction, the analogous specification must be the quality of the odorant used and its concentration. It is assumed that an alcohol and an aldehyde may be compared for odor intensity as easily as red and blue lights can be matched for brightness. In this paper we describe the construction and use of a 'standard' scale of olfactory intensities.

Materials. The olfactory intensity of a given odorant can most easily be changed by altering its concentration. Although the relation between concentration and intensity may not be simple, the physical characteristics of odorants (vapor pressure, thermodynamic activity, etc.) may be computed from concentration and for this reason concentration units are employed. The technical difficulties encountered with the use of gases in olfactory experiments make liquids superior for our purpose and lead to consideration of the selection of a diluent and 'standard' odorant.

The properties required of a diluent suitable for work in olfaction are the following: it should be odorless; a good organic solvent; readily obtainable; and chemically stable. The most suitable substance we have found thus far is benzyl benzoate. It has the low Crocker-Henderson rating 3111 and its odor is not readily detected by most *O*s.[8] Moreover, its low vapor pressure provides a less saturated vapor for inhalation and minimizes the problem of evaporation.

A suitable standard odorant should be chemically stable and have a strong, characteristic, but not unpleasant, odor. A low vapor pressure is important here, since very low concentrations must occasionally be used and evaporation can be a serious problem. If the *O*s are able to see the solution they are smelling, the 'standard' should preferably be colorless. Mixtures have not been considered because the physico-chemical properties as well as the quality of the odor of complex mixtures are difficult to specify. Among the several substances we have studied n-heptanal comes closest to meeting the above requirements.[9]

Method. The manner in which the odorant is delivered to the olfactory end organ constitutes the first problem. Techniques involving tubes in the nose and the injection of odorants under pressure[10] were rejected in this study for reasons of discomfort and complexity and the difficulty of interpreting pressure as an olfactory stimulus.[11] Since human *O*s report odors[12] and electrical discharges from

[8] E. C. Crocker and F. M. Dillon, Odor directory, *Amer. Perfumer*, 53, 1949, 297-306.

[9] L. H. Beck, Lawrence Kruger and Paul Calabresi, Observations on olfactory intensity: I. Training procedure, methods, and data for two aliphatic homologous series, *Ann. N. Y. Acad. Sci.* 58, 1954, 225-238.

[10] C. A. Elsberg and Irvin Levy, A new and simple method of quantitative olfactometry, *Bull. Neurol. Inst. N. Y.*, 4, 1935, 5-19.

[11] F. N. Jones, A test of the validity of the Elsberg method of olfactometry, this JOURNAL, 66, 1953, 81-85.

[12] B. M. Wenzel, Differential sensitivity in olfaction, *J. Exper. Psychol.*, 39, 1949, 129-143.

the olfactory bulb occur in animals when 'odorless' air is injected under pressure,[13] techniques employing blast injection introduce extraneous problems. Natural sniffing, on the other hand, suffers the disadvantage of wide variability, which may, however, be reduced by suitable training. *O*s rapidly learn that it is necessary to sniff in the same fashion each time if the odor sensation is to remain constant. We have used a natural technique with trained *O*s sniffing from test tubes containing the odorants.

As a routine procedure we employed a dilution series of an odorant to train the *O*s. Each *O* was required to learn to arrange a randomly presented set of sample tubes in order of concentration. Training was continued until the *O* was able to order the tubes correctly in two consecutive sessions before the collection of data was begun. During this period *O* learned to employ a constant sniff. The

TABLE I

OLFACTORY SCALE UNITS

Tube	n-heptanal		
	volume (%)	Mole fraction	Vapor pressure*
1	100	1.0000	4.129
2	50	0.5852	2.416
3	25	0.3198	1.321
4	12.50	0.1677	0.692
5	6.25	0.0860	0.355
6	3.12	0.0435	0.180
7	1.56	0.0219	0.090
8	0.78	0.0110	0.045
9	0.39	0.0055	0.023

* At 20° C. in mm. Hg.

sniff, of course, varied from *O* to *O*, but since a similar sniff is used in sampling both the test and the standard odorants, presentations should be equivalent.

In every case, 1 ml. of the odorant was placed in a 10 × 75 mm. test tube, kept corked when not in use, from which the *O* sniffed. The 'standard' scale consisted of a geometric dilution series of n-heptanal in benzyl benzoate, constructed so that tube 1 was 100% n-heptanal and tube 'n' was $100\%/2^{n-1}$ n-heptanal in benzyl benzoate by volume. Table I describes the properties of the standard solutions in this scale. The usual range of the scale may stop before or extend further than Tube 9, which is near threshold for most *O*s tested to date. The mole fractions were calculated using density values of 0.850 (20/4) for n-heptanal, and 1.120 (20/4) for benzyl benzoate. The vapor pressure of this diluent (benzyl benzoate) at 20° C. was measured and found to be approximately 0.17 mm.

After familiarizing himself with the intensity of the n-heptanal standard solutions, *O* was presented with 1 ml. of an unknown odorant in a corked 10 × 75 mm. test tube. After removing the cork, sniffing, and replacing the cork, *O* selected a tube from the standard scale, removed the cork, and sniffed. Following a short

[13] E. D. Adrian, The electrical activity of the mammalian olfactory bulb, *E. E. G. Clin. Neurophysiol.*, 2, 1950, 377-388.

interval he repeated the procedure, moving in the appropriate direction on the scale. After several repetitions, *O* usually had reached a decision about the 'position' of the unknown on the scale and registered it accordingly.

Evaluation of method. Without great difficulty most *Os* could match the unknowns (excepting organic liquids of high vapor pressure) to the standards, or place the unknown between two standards, in which case 0.5 was added to the lower tube number in expressing the *O*'s judgment. *Os* with more acute sensitivity were often able, especially at the stronger end of the scale, to rate odorants by tube number, as for example 2¼, 2½, 2¾, etc.; and, in fact, such interpolations were encouraged since most people are unaware of the full extent of their olfactory sensitivity.

The rapidity of olfactory adaptation precluded the rapid collection of data. Frequent rests for recovery are necessary. The effects of adaptation may be minimized and the whole experiment may proceed more rapidly by having *O* clear his nasal passages by short, forceful expirations. A room well ventilated with fresh air is also desirable.

Care was taken to avoid conducting tests in the presence of strong extraneous odors, although no attempt was made to approximate an 'odor-free' environment. The non-transient odors existing in the testing room, presumably affecting the odor of the unknown and the standard equally, were ignored.

Using this procedure we have rated the olfactory intensity of numerous aliphatic and aromatic compounds with widely different physico-chemical properties, finding that they fall within this scale. Among the substances we have studied, only the very volatile organic liquids (nasal irritants which are usually dangerous to inhale) led to inconsistent, statistically meaningless data, presumably because these substances could not be uniformly and repeatedly sniffed by the *Os*.

To evaluate the internal consistency of the scale, four other dilution scales were constructed for intercomparison. Preliminary results showed that over a wide range two different solutions (aliphatic alcohols) of differing concentrations which were matched to be equal to each other in intensity, were also matched to approximately the same point on the n-heptanal standard scale.[14]

An indication of the sensitivity of different parts of the scale can be obtained from Fig. 1. The ordinate is a measure of dispersion for those intensity values lying in the half-interval ($n \leq x < n.5$, $n = 1, 1.5, 2.0,$

[14] Beck *et al., op. cit.*, 235.

2.5. . . . . 8.0. Below each point is the number of experiments considered in the half-interval. Each experiment consisted of 8 or 10 judgments on one unknown by one *O*. A linear estimate of the standard deviation with efficiency not less than 96% was computed for each experiment, the experiments grouped into half-intervals by intensity, and the mean of the standard deviations computed for each half interval.[15] The reliability of the statistic is proportional to the number of experiments in the half-interval. In a few cases, the same unknown was used in different experi-

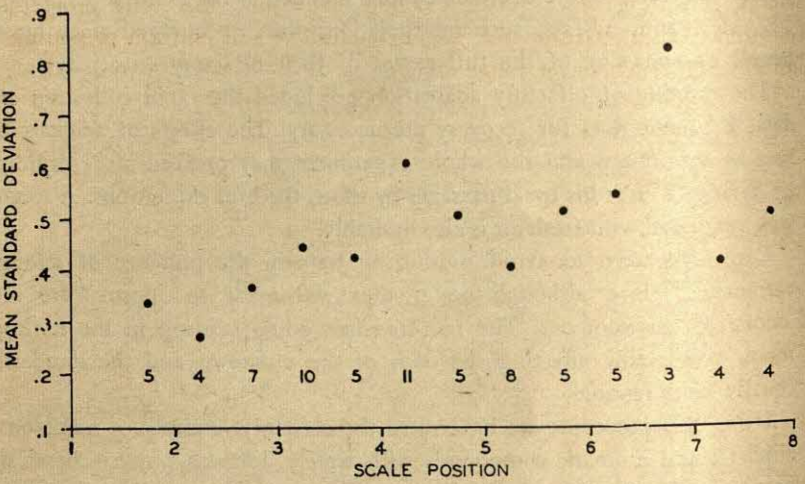


FIG. 1. COMPARISON OF SENSITIVITY WITH SCALE-POSITION

ments, but no *O* performed two experiments with the same unknown. In all, 652 sniffing judgments and 66 unknowns are represented in this graph. A high index of dispersion in one section of the scale would indicate that *O* is unable to make consistent judgments in that intensity area, *i.e.* the scale demands finer discrimination than *O* is capable of making, and that the intervals should, therefore, be increased. Conversely, a low dispersion would indicate that the scale is too crude and more standard positions should be inserted. An ideal scale would, of course, produce the same measure of statistical variation throughout its range. Fig. 1 shows that the scale is not grossly skew in this respect. The statistical variation at the more intense end is somewhat less than elsewhere, which implies

[15] W. J. Dixon and F. J. Massey, Jr., *Introduction to Statistical Analysis,* 1951. Chap. 16.

that more standard positions could profitably be inserted with a probable slight gain in discrimination.

SUMMARY

The use of matching of olfactory intensities, analogous to the procedures used in vision and audition, is suggested as a means of measuring supra-threshold olfactory intensity. A scale based on various concentrations of n-heptanal diluted in benzyl benzoate is described, and its appropriate physico-chemical units are presented. Preliminary results on several substances, including the aliphatic alcohols, show that this scale provides consistency of comparison and is a fairly sensitive and uniform measuring instrument throughout the range studied.

FOURIER ANALYSES FOR CURVES OF AFFECTIVE VALUE OF COLOR AS FUNCTIONS OF HUE

By JOHN S. STAMM, California Institute of Technology

In the field of quantitative psychology methods of mathematical analysis have frequently been applied to the measurement and interpretation of psychological responses. This type of analysis may be relevant to investigations of color preferences. Guilford has shown that the psychological variables of affective value of colors and their visual properties of hue, tint, and chroma are related to another by continuous functions.[1] He employed the Munsell system of color notation and a rating scale of affective responses for measuring these continua.

Since hue, according to the Munsell notation, can be expressed in terms of a circular function, its relationship to affective value (*AV*) may be expressed mathematically by Fourier equations. In his earlier investigation, using 40 different color samples, Guilford subjected the empirical curves of *AV* as function of hue to Fourier analyses.[2] He found that "two components, the first and third harmonics, account almost entirely for the variations in affective value as dependent upon hue."[3] In a more extensive investigation he selected a total of 316 different color samples of 12 different hues, according to the Munsell color notation.[4] The *AV*s of these samples were rated by groups of 20 men and 20 women and the medians of the group-judgments were computed. Curves were then constructed showing the *AV*s, for each sex, as functions of hue, for constant designations of tint and chroma. These curves were generally continuous and exhibited several marked regularities. The small differences between corresponding curves for men and women consisted of slightly higher *AV*s for women and somewhat greater fluctuations of their curves of *AV*. It was therefore decided to analyze curves for women, since regularities found for this group should also apply to the men's responses. Mathematical analyses of these curves might point out certain inherent relationships between affective judgments and the visual dimensions of color.

Materials. Guilford's color samples were selected on the basis of their distribution in the Munsell *Book of Color*.[5] In this notation hue attributes are found on the scale of a 'color circle.' This scale was constructed by spacing the five principal hues—red, yellow, green, blue and purple—at equal intervals on the circle and then

* Accepted for publication February 1, 1954. This investigation was conducted at the University of Southern California. The author is indebted to Professor J. P. Guilford for supplying the data and for supervising this study.

[1] J. P. Guilford, The affective value of color as a function of hue, tint, and chroma, *J. Exper. Psychol.*, 17, 1934, 342-370; A study in psychodynamics, *Psychometrica*, 4, 1939, 1-23.

[2] Guilford, 1934, *op. cit.*, 354.

[3] Guilford, 1939, *op. cit.*, 4.

[4] Guilford, 1939, *op. cit.*, 5.

[5] A. H. Munsell, *Book of Color*, 1929.

placing intermediate hues between them.[6] Chroma, a measure of the 'strength of color,' is expressed according to a 10-interval scale, the maximal saturation being 10. Munsell defined 'value' as "the quality by which we distinguish a light color from a dark one."[7] To avoid confusion in terminology, Guilford called this dimension 'tint.' This scale is divided into 10 equal steps, zero being equivalent to pure black and 10 equivalent to pure white.

The judgments of *AV* of the color samples were expressed in terms of a rating scale of 11 steps, zero being labeled as "the most unpleasant imaginable," 10 as

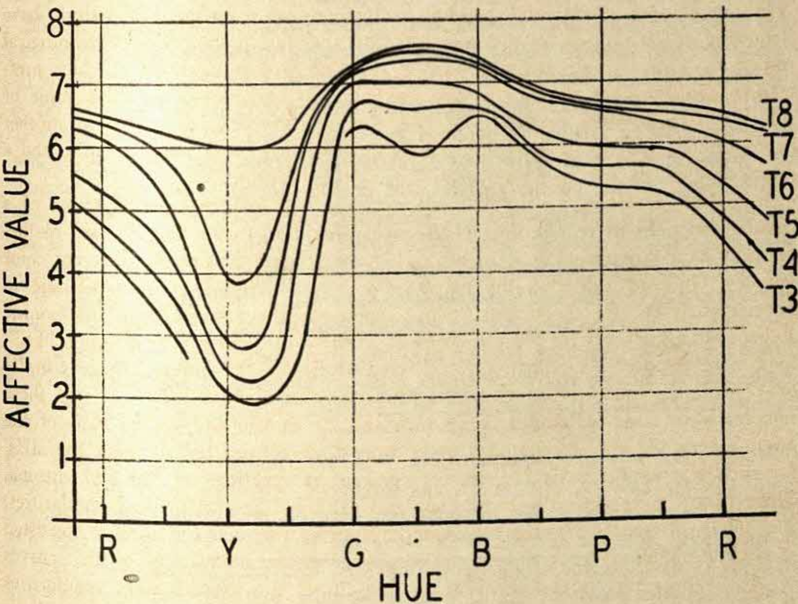


FIG. 1. AFFECTIVE VALUE OF COLOR AS FUNCTION OF HUE
The curves are based on the Median judgments of 20 *Ss* (women). Chroma is 6 for all curves.

"the most pleasant imaginable," and 5 as "indifferent."[8] The reliability of this scale, determined for 20 *Ss* (men), was found to be 0.94.

The results of Guilford's investigation, using 316 color samples, were presented in a series of graphs which show the median *AV*s as functions of hues.[9] Fig. 1 illustrates a typical set of curves for six designations of tint, when the chroma was held constant at C6. In the present investigation two series of curves were subjected to Fourier analyses, one for constant chroma (C6), at approximately average saturation, and one for constant tint (T5), equivalent to neutral gray. Since certain of the curves with extreme notations of tint or chroma are not continuous, they could

---

[6] T. M. Cleland, *A Practical Description of the Munsell Color System*, 1931.
[7] Cleland, *loc. cit.*, 7.
[8] Guilford, 1939, *loc. cit.*
[9] Guilford, 1939, *loc. cit.*

not be readily analyzed. The present analysis, covering the range of tint between T4 and T8, and that of chroma between C2 and C8, includes therefore all the tints and chromas available for the total range of hue.

Fourier analysis. An empirical curve which is continuous and periodic may be expressed by an equation consisting of a finite number of simple sine and cosine terms. The computational problem consists in the determination of these Fourier coefficients from the knowledge of a number of equally spaced points on the empirical curve. Since each coefficient is composed of the sum of a number of sine or cosine terms, the computational process would be a formidable task.[10]. Fortunately, this task can be simplified by following worksheets which have been prepared for analyses based on 24, 12, and 6 points, respectively, taken from the empirical curve. The procedures suggested by Manley were followed in the present computations.[11] The number of unique Fourier coefficients which may be computed for a given curve is equal to the number of empirical points used in the calculations.

In carrying out the Fourier analyses, red on the Munsell notation was taken as zero degrees of the circular functions. The analyses of the curves with Tint 5 were based on the 24 coördinate scheme which yielded 12 cosine, 11 sine, and one constant term. Since the resulting equations had zero magnitudes for all harmonics higher than the seventh, and generally also for the 5th and 6th harmonics, the 12 coördinate scheme was used for the computations of the remaining curves with chroma of 6. The resulting Fourier equations were then converted to terms containing sine coefficients and phase angles (rather than cosine terms), which reduced the number of terms in the equations.

Results: (1) Fourier coefficients. Fig. 2 illustrates the relation between one empirical curve (Tint 5, Chroma 6), and the first 4 harmonic components obtained from the Fourier analysis. The instantaneous sums of these components are shown by the circled dots for several points along the hue scale. On the basis of this graphical construction the original empirical curve and the analytical curve, derived from the Fourier analysis, appear to coincide quite well. The final equation for this function may be expressed by:

$$y = 5.50 - 1.65 \sin (x + 38°) - 1.08 \sin 2 (x - 32°) + 0.50 \sin 3 (x + 7°)$$
$$- 0.32 \sin 4 (x + 22°) - 0.17 \sin 5 (x - 4°)$$

The amplitudes and the phase angles of the sine coefficients for the computed Fourier equations are listed in Tables I and II, respectively. The zero components in Table I indicate the magnitudes of the constant term, while the other component designations correspond to the harmonic terms in the equations. The accuracy of the Fourier equations was estimated by constructing curves similar to the ones in

[10] The equation for computing the cosine coefficient of the k harmonic is:

$$a_k = \frac{2}{N} \cdot \sum_{r=1}^{N} Y_r \cos i\, x_r,$$

where N is the number of points taken from the empirical curve, Y_r the ordinate, and x_r the abscissa of the point r. A similar equation may be found for computing sine coefficients.

[11] R. G. Manley, *Waveform Analysis*, 1945, 183-202.

Fig. 2. Corresponding empirical and analytical curves coincided closely, according to inspection of these figures.

(2) Interpretation of Fourier equations. The Fourier analyses never yielded harmonics higher than the seventh, although the magnitudes for 12 possible components were computed. Moreover, the magnitudes of the 6th and 7th harmonics were never larger than 0.15, which may correspond to the experimental and graphical errors. Amplitudes of successive harmonic components in each equation also generally decreased, the fifth having the smallest magnitude.

The first and second harmonic components together account for the major portions of the harmonic fluctuations. Their contributions may be expressed by the

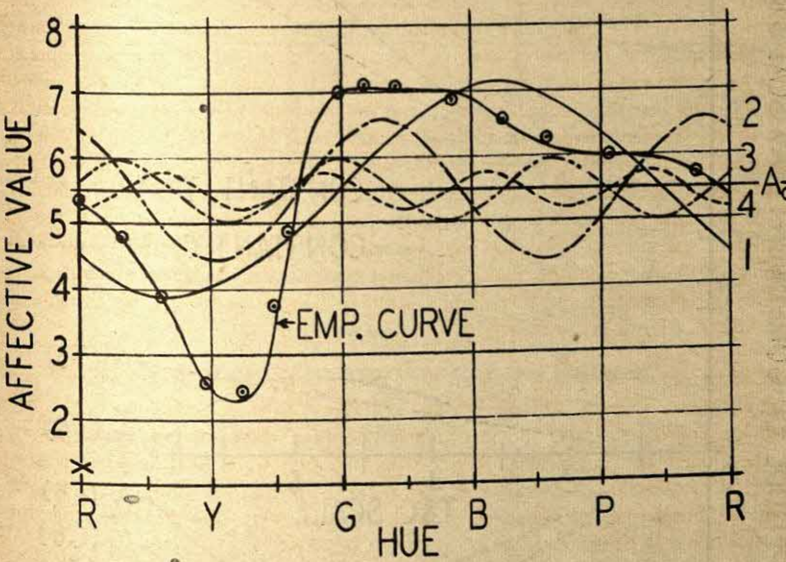


FIG. 2. HARMONIC COMPONENTS FOR CURVE OF AFFECTIVE VALUE OF COLOR AS FUNCTION OF HUE, FOR WOMEN, AT CONSTANT CHROMA 6 AND TINT 5

The order of the harmonics is indicated along the right margin of the figure. The empirical curve is shown by the heavy line. Instantaneous points obtained from the Fourier equation are shown by the circled dots. The constant term of the equation is 5.5.

ratio of the sum of the magnitudes of the first and second components to the sum of the magnitudes of all harmonic components. For the 8 computed equations, this ratio ranged between 0.68 (chroma 6, tint 6) and 0.92 (chroma 6, tint 8), the average being 0.76. The first and second harmonic components together therefore contribute on the average about 75% of the total fluctuations of the curves of affective value.

An examination of Table I indicates certain variations in the magnitude of the Fourier coefficients for increasing chroma and tint designations. These functions, for the constant terms and the first three harmonics, are shown graphically in Fig. 3. According to this figure the magnitudes of the constant terms (A_0) are related

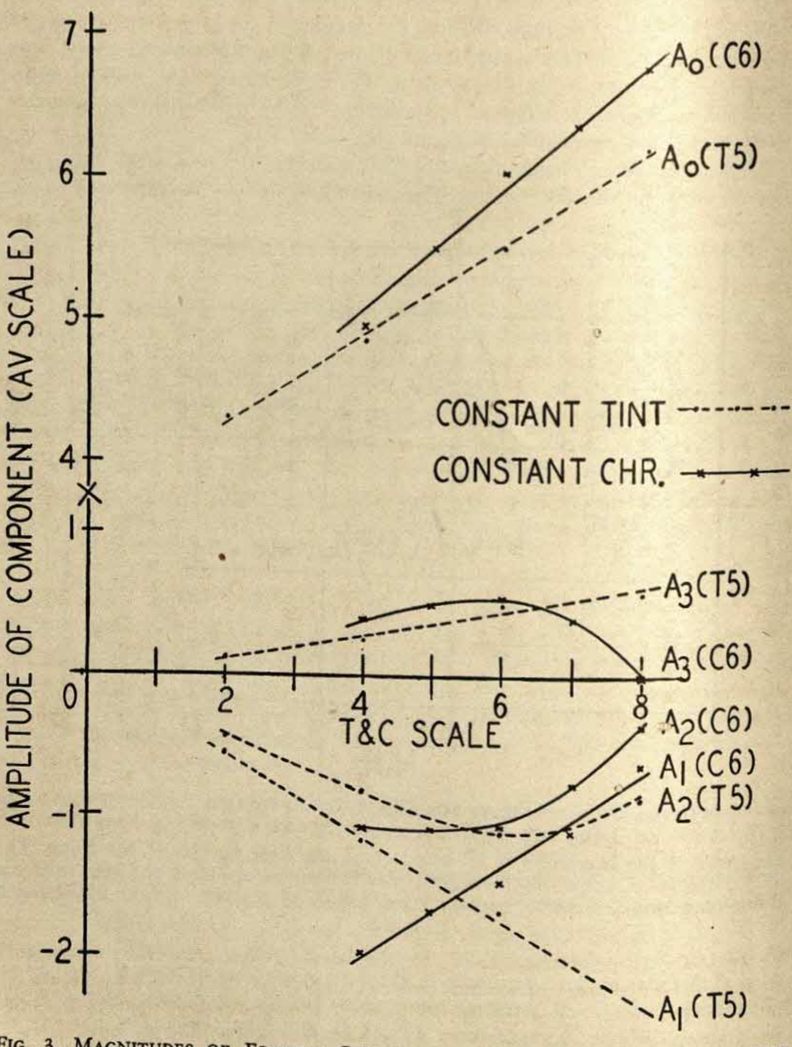


FIG. 3. MAGNITUDES OF FOURIER COMPONENTS AS FUNCTIONS OF TINT (WITH CONSTANT CHROMA 6) AND CHROMA (WITH CONSTANT TINT 5)

The abscissa is calibrated in terms of the Munsell notation for tint and chroma; the ordinate in terms of the scale of affective value. Curves are shown for the constant terms (A_0) and the first three harmonics (A_1, A_2, and A_3).

approximately linearly to both tint and chroma designations. The coefficients of the first harmonic term also change approximately linearly as functions of tint and chroma. The relationships between the coefficients of higher harmonics and tint and chroma, respectively, are somewhat more complex, but, on the basis of the curves in Fig. 3, these functions may be expressed by first or second order equations.

According to Table II the phase angles for the first and second harmonics, are within relatively narrow limits, but are more widely distributed for the higher

TABLE I

MAGNITUDES OF HARMONIC COMPONENTS FOR CURVES OF AFFECTIVE VALUE OF COLOR

Notation for		Magnitude of harmonic component*							
chroma	tint	0	1	2	3	4	5	6	7
2	5	4.32	−0.55	−0.45	0.12	−0.17	0	0	0
4	5	4.86	−1.17	−0.80	0.26	−0.44	0.13	−0.13	−0.10
6	5	5.50	−1.65	−1.08	0.50	−0.32	−0.17	0	0
8	5	6.18	−2.46	−0.85	0.58	−0.41	0.19	0.10	0
6	4	4.95	−1.94	−1.08	0.40	−0.28	0.22	—	—
6	6	6.02	−1.46	−1.05	0.55	0.39	−0.20	—	—
6	7	6.35	−1.09	−0.76	0.40	0.25	−0.13	—	—
6	8	6.75	−0.61	−0.36	0	0	0	—	—

* According to Guilford's scale of affective value.

TABLE II

PHASE ANGLES OF HARMONIC COMPONENTS FOR CURVES OF AFFECTIVE VALUE OF COLOR

Notation for		Phase angle (degree) for harmonic component*						
chroma	tint	1	2	3	4	5	6	7
2	5	50	−22	30	7	0	0	0
4	5	43	−29	0	22	18	15	0
6	5	38	−32	7	22	− 4	0	0
8	5	31	−33	8	27	0	15	0
6	4	30	−38	− 6	10	13	—	—
6	6	35	−33	10	−12	6	—	—
6	7	37	−36	6	−17	0	—	—
6	8	64	−40	0	0	0	—	—

* Red located at 0°.

harmonics. The range of phase angles for the first harmonic is 19° for the constant tint, and 8° for the constant chroma series (exclusive of the curve for Chroma 6, Tint 8, which shows unusually small fluctuations from its mean *AV*). For the second harmonic the corresponding ranges are 11° for the constant tint, and 8° for the constant chroma series, respectively.

These restricted ranges of phase angles are illustrated in Fig. 4A, which shows the spread of the maximal and minimal points for the first and second harmonic components on the scale of hues. This figure indicates that, for the total range of chroma and tint used, these two harmonic components are placed within narrow limits on the hue scale.

Discussion. The present analysis indicates a number of orderly relations between *AV* of color and the visual dimensions of hue, tint, and chroma: (a) the constant terms of the equations increase linearly as functions of both tint and chroma; (b) the first two harmonics account for approximately 75% of the fluctuations of the curves of *AV* and no appreciable components higher than the fifth harmonic have been obtained; (c) the magnitudes of the first and second harmonics are, moreover, relatively simple functions of both tint and chroma; and (d) the maximal and minimal points of the first two harmonics fall within relatively narrow limits on the hue scale.

This analysis suggests the possibility of developing a 'master equation' for expressing *AV*s as functions of the three visual dimensions. By computing empirical equations for the curves of the constant terms and the first two harmonics in Fig. 3, it should be possible to express the magnitudes of these components as functions of tint and chroma. Similar equations may be developed for the phase angles of the first two harmonics as functions of tint and chroma. Such a 'master equation,' would, however, appear rather complex and may not reveal new relationships between *AV*s and the psychophysical dimensions of color.

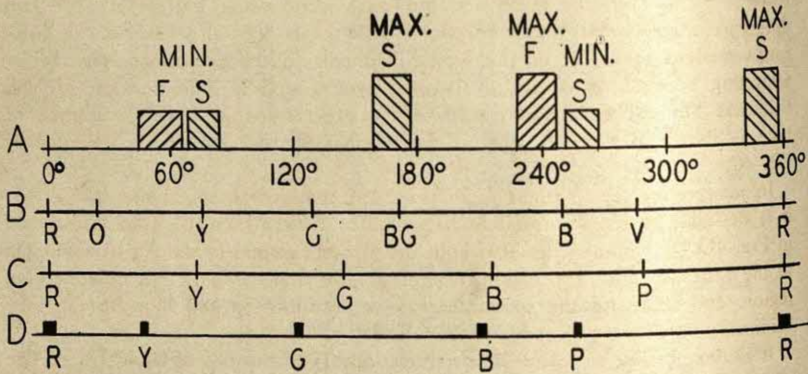


FIG. 4. SCALE OF HUES

(A) Location of maximal and minimal points for the first (F) and second (S) harmonics of the analyzed curves—scale is in degrees with red at zero. (B) Designation of hues employed in Guilford's analysis. (C) Designation of hues according to the original Munsell notation. (D) Range of Munsell hues for various tints and chromas.

One limitation in developing such an equation at the present time lies in certain inconsistencies of the Munsell notation, on which the present color samples were based. Thorough investigations concerning the spacings of Munsell colors have in recent years been carried out by Newhall, Nickerson, and Judd, who related the Munsell designations to the physical dimensions of color.[12] They found that, to a

[12] S. M. Newhall, D. Nickerson, and D. B. Judd, Final report of the O.S.A. subcommittee on the spacing of the Munsell colors, *J. Opt. Soc. Amer.*, 33, 1943, 385-418.

first approximation, "Munsell hue, value, and chroma may be taken as the correlates of hue, lightness, and saturation, respectively, of object colors."[13] The relationships between corresponding functions were found to be, however, not linear, but of an exponential nature. These investigators also re-defined the Munsell colors according to accurate measurements of psychophysical judgments. To compute more correct mathematical functions between *AV*s and hue, tint, and chroma, it would therefore be necessary to designate the Munsell color samples according to the notation recommended by these investigators.

The earlier investigation by Guilford pointed to the importance of the first and third harmonics in accounting for the fluctuations of the curves of *AV* of color. The curves which Guilford subjected to Fourier analyses were only in general agreement with those analyzed in the present study. The former curves showed marked peaks of *AV* in the blue region whereas the present series of curves showed flatter tops between the green and blue-purple regions of hue (Fig. 1).[14] This difference may account largely for the importance of the third harmonic in Guilford's analysis, and that of the second harmonic in the present analysis. In interpreting his findings, Guilford also employed a slightly different scale of hues than the Munsell designation, used in the present analysis. His designation of hues is shown in Fig. 4B, whereas the original Munsell notation is indicated by Fig. 4C.

In interpreting the locations of maximal and minimal points, shown in Fig. 4A, it is desirable to use the revisions suggested by Newhall, et al., which are shown in Fig. 4D.[15] It appears then that both the first and second harmonics have minima in the yellow region. The maximal points for the first harmonic fall near the blue region, and those for the second harmonics near the red and blue-green points. Minimal points for the second harmonic also fall near purple.

It is not possible to relate these findings readily to theories of color vision. The present analyses emphasize however, the importance of the regularity in color preference as functions of the three dimensions of color: hue, tint, and chroma.

SUMMARY

The present analysis is concerned with a series of curves of affective value (*AV*) of color as functions of hue, for constant tint and chroma designations, according to the Munsell color notation. The curves were constructed by Guilford, as the result of judgments for 316 color samples by a group of 20 women. The series of curves includes the tints and chroma available for the total range of hue. When these curves were subjected to Fourier analyses the following results were obtained.

(1) The constant terms of the Fourier equations increase linearly as functions of both tint and chroma.

(2) The first two harmonic components account for approximately 75% of the fluctuations of the curves of *AV*. No appreciable components higher than the fifth harmonic were obtained.

[13] Judd, Basic correlates of the visual stimulus, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 852.

[14] Guilford, 1934, *op. cit.*, 356.

[15] Newhall, et al., *op. cit.*, based on Fig. 13, p. 415.

(3) The magnitudes of the first two harmonics are related to tint and chroma by first or second order equations.

(4) The maximal and minimal points of the first and second harmonics fall within relatively narrow limits on the hue scale.

The accuracy of these findings may be limited by inconsistencies in the spacings of the Munsell notation. More correct designations have recently been developed and these should be applied to further investigations on color preference as a function of the visual dimensions.

FIGURAL AFTER-EFFECTS WITH COLORED STIMULI

By JULIAN E. HOCHBERG, Cornell University, and WILLIAM TRIEBEL, New York Hospital, Westchester Division

Köhler and Wallach[1] have explained Gibson's tilted-line effects[2] and other phenomena, in terms of a physiological model which has far-reaching perceptual implications beyond the figural after-effects themselves. Although the model was developed for the achromatic colors—black, grays, white—it would seem worthwhile to establish its limits of applicability, and, possibly, to aid understanding of the chromatic processes, by determining to what extent the 'contour-repulsion' figural after-effects are related to *chromatic* color.

EXPERIMENT I

Procedure. The stimulus-figures used in Experiment I differed from their backgrounds in chroma and brightness. The procedure involved seven stages as follows:

(*1*) Fore-test: presentation of each of the three test-figures (T_a, T_b, T_c) at a distance of 2 ft.
(*2*) Stimulation: satiation with an inspection-figure (*I*-figure) at the same distance for 3 min.
(*3*) After-test: with T_a-figure.
(*4*) Satiation: with the *I*-figure for 60 sec.
(*5*) After-test: with T_b-figure.
(*6*) Satiation: with the *I*-figure for 60 sec.
(*7*) After-test: with T_c-figure.

Instructions. For the odd-numbered stages (*1, 3, 5,* and 7), the instructions to the *Ss* were: "Please keep your eyes on the small fixation-point. While doing this, tell me if one circle appears smaller than the other." For the even-numbered stages (*2, 4,* and *6*), the instructions were: "Please keep your eyes on the fixation-point; do not move your head or eyes—you may blink however."

The *I*-figures consisted of two annuli—a red (Munsell No. 5) and a green (Munsell No. 5)—each 1 mm. thick and 5 mm. in radius. The center of each was 12 mm. from the fixation-point, the red annulus to the right and the green annulus to the left. The background was white.

The *Os* (10 in number) were divided into two groups of 5 each. The *T*-figures (*a*, *b*, and *c*) for Group I were respectively red, green, and black pairs of annuli; one pair of each color on a white background. The annuli were 0.5 mm. thick and 4 mm. in radius. They were placed each side of the fixation-point, concentric with

* Accepted for publication March 29, 1954.

[1] Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Philosoph. Soc.*, 88, 1944, 269-357.

[2] J. J. Gibson, Adaptation, after-effect, and contrast in the perception of curved lines, *J. Exper. Psychol.*, 16, 1933, 1-31.

the previously presented annuli of the *I*-figure. Thus the *T*-figures fell within the areas satiated by the *I*-figures and they should therefore appear smaller in consequence of the figural after-effect of 'contour repulsion.' The *O*s of Group II were presented with the *T*-figures in reverse order.

Results. The effects of satiation with the red and the green *I*-figures were compared for different size-effects on the three colors of the *T*-figures. No consistent

TABLE I

RESULTS OF EXPERIMENT I, II, AND III

	No. of Ss	Condition	T-fig.	No. of Ss reporting: right larger	left larger	equal
Exper. I	5 (Group I)	fore-test	a	0	0	5
			b	0	0	5
			c	0	2	3
		after-test	a	0	0	5
			b	0	0	5
			c	0	3	2
	5 (Group II)	fore-test	a	0	0	5
			b	0	0	5
			c	0	0	5
		after-test	a	0	2	3
			b	0	1	4
			c	0	0	5
Exper. II	10	fore-test	a	0	2	8
			b	0	0	10
			c	0	2	8
		after-test	a	0	9	1
			b	0	9	1
			c	0	10	0
Exper. III	10	fore-test	a	0	0	10
			b	0	0	10
		after-test	a	0	0	10
			b	0	0	10

differences ascribable to the chroma of the *I*- or *T*-figures were obtained (see Table I). Either these stimulus-figures had no figural after-effects or, if they did, they were equal for the annuli of the two *I*-figures and independent of chroma. These alternatives were separated in Experiment II.

EXPERIMENT II

Whether the lack of differences in Experiment I occurred through the independence of figural after-effects and chroma, or through a complete absence of figural after-effect due to some procedural artifact, was unknown. We therefore sought information in Experiment II upon this point.

Procedure. Ten new *O*s followed the procedure of Experiment I, except that only

one annulus (red) was used as the *I*-figure. This was placed to the right of the fixation-point.

Results. The annulus on the right of each T-figure appeared smaller than that on the left for most of the *O*s (see Table I), *i.e.* figural after-effects were obtained. The results of Experiment I thus were most probably due to the inability of chromatic differences to cause differences in the amount of the after-effect, rather than to an absence of the after-effects themselves. Is chroma then completely irrelevant to figural satiation?

Experiment III

Since difference in chroma has not caused difference in after-effect, we sought in Experiment III to determine whether after-effects can be obtained from an *I*-figure which differs from its ground in terms of chroma and little, or not at all, in brightness.

Procedure. The *I*-figure was identical to that of Experiment II but with a gray background (Munsell No. 5), hence with little (if any) brightness difference between figure and ground. The *T*-figures were presented to 10 new *O*s: T_a, after a satiation-time of 3 min.; T_b after a second satiation-time of 2 min.

Results. No consistent size-difference obtained between the test-annuli which fell in the region previously satiated by the *I*-figures, and the other test-annuli (*i.e.* no measurable figural after-effect was obtained).

Summary and Conclusions

When figures differ from their ground in terms of both chroma and brightness, figural after-effects are obtained which do not, however, vary with change in chroma of *I*- or *T*-figure. When figures differ from their ground largely (or only) in chroma, no measurable figural after-effects are obtained. The present conclusions are that *chromatic differences can neither modify nor generate the phenomenal after-effect.* This poses something of a problem. Perhaps brightness differences are necessary (and chromatic differences insufficient) for those figural processes presumably responsible for the after-affects. (The figures were near the Liebman level,[3] although their form was perfectly clear. If form can be perceived independent of figural process inducing satiation, then the Köhler-Wallach physiological model—which is isomorphic to experience—will require revision. It may be, however, that figural after-effects were occurring which were not measurable by the usual techniques. Perhaps the differential satiation of portions of a reversible figure, which has been employed elsewhere,[4] would provide a more sensitive method.

[3] S. Liebmann, Ueber das Verhalten farbiger Formen bei Helligkeitsgleichheit von Figur und Grund, *Psych. Forsch.*, 9, 1927, 300-353.

[4] J. E. Hochberg, Figure-ground reversal as a function of visual satiation, *J. Exper. Psychol.*, 40, 1950, 682-686.

RANDOMIZATION OF A BINARY SERIES

By BRUCE M. ROSS, Lawrence College

The hypothesis investigated in the present study is the concept widely held that man is incapable of randomizing a series of symbols. Specifically, it is assumed that naïve persons, when directed to construct a random series, will alternate sequential symbols more frequently than chance; that they will not allow sufficiently for the clusters of similar symbols which do occur in any random series with limited alternatives.[1]

In the present experiment our *Ss* were required to construct three binary series of 100 symbols in length and of frequencies 50-50, 60-40, and 70-30. A count of the number of unlike alternations was made for every series and compared with the mean expected for a mathematically random series of the same frequency to determine whether an hypothesis of superfluous alternations was tenable. An analysis of variance was also performed to determine whether symbol-frequency and presentation-sequence were significant variables.

Subjects. The Ss, 60 in number (40 women and 20 men), were undergraduate students in elementary psychology at the University of Wisconsin. The experiment was administered individually and in groups of two or three. When administered in groups, the *Ss* were separated by partitions and directed to follow different experimental sequences.

Procedure. S was presented with two open card files (3″ × 5″), each filled with an unknown number of cards; an empty card file, which was placed between the two with cards; a stamp pad; and two rubber stamps marking *X* and *O* respectively. After these materials were properly placed and *S* was comfortably seated, *E* read the following directions to him.

This experiment is *not* a timed test. [If groups of *Ss* were used, the following was added.] Neither is it a social nor group experiment. The reason you are here together is solely for convenience in administration.

The cards before you are to be placed in what you consider to be a *chance* or *random* order, by which is meant a sequence such as might be obtained by rolling a die or flipping a coin. All the cards in the file by the *X*-stamp are to be stamped with it and similarly those in the file by the *O*-stamp are to be stamped with it. Remove a card from either file, then stamp it and place it on the empty file in the middle. Try and stamp the card so the symbol *X* and *O* is straight and well-centered. Do *not* remove a card once placed in the middle file. Remember, the middle file contains your randomized sequence which will be scored to determine how well you succeeded in this experiment. When you are done with the first two files, I will give you two more files of cards to randomize. The experiment will be concluded by your filling out a questionary. Are there any questions? [If so, they were immediately answered.]

---

* Accepted for publication March 22, 1954, from the Department of Psychology, University of Wisconsin. This study was directed by Professor David A. Grant.

[1] Hans Reichenbach, *The Theory of Probability*, 1949, 152-153.

S then stamped and filed the cards at his own desired speed. As soon as he exhausted his supply, he was handed two more files of cards, and again when these cards had all been stamped and filed he was given two additional files of cards. Both rubber stamps and deck-frequencies were randomized with regard to right and left positions. The answers to the questionary revealed that the *S*s were not trained in the theory of probability, and that none of them approached the task from a statistical point of view.

Experimental design. The six possible sequential permutations of three experimental conditions were utilized in the study. The *S*s were assigned to sequences randomly. The experimental conditions were: (1) 50 cards of each of two symbols, *X* and *O*; (2) 60 cards of one symbol and 40 cards of the other; and (3) 70 cards

TABLE I

EXPERIMENTAL SEQUENCES

1	2	3	4	5	6
60–40	70–30	50–50	70–30	50–50	60–40
50–50	60–40	70–30	50–50	60–40	70–30
70–30	50–50	60–40	60–40	70–30	50–50

of one symbol and 30 of the other. No comments were made concerning the total number of cards or the disparity between number of cards in each file under Conditions (2) and (3). Care was taken to see that each card was individually stamped and placed in the center file. One *S* was dropped from the study for failing to follow this procedure. Although the experimental conditions imposed exact frequency-limits, it was anticipated that the use of frequencies other than 50-50 would permit us to generalize from our results.

The experimental design is summarized in Table I. Each sequence had 10 *S*s.

Results. The data were summarized by Wald-Wolfowitz's method.[2] The mean number of alternations of unlike symbols expected for any frequency was calculated by the formula $[2mn/(m+n)]+1$, where m and n are the two alternative sequential events. The theoretical mean number of alternations obtained by this formula for each sequence of 100 symbols is: 51 for a 50–50 sequence, 49 for a 60–40 sequence, and 43 for a 70–30 sequence.

The 60 *S*s yielded mean values of 48.65 for a 50–50 sequence, 45.85 for a 60–40 sequence, and 40.37 for a 70–30 sequence. Clustering is thus, on the average, slightly greater than is predicted mathematically for each of the three conditions, a contradiction of the prediction that the opposite trend of over numerous alternations would be obtained. Median alternation-scores close to the values of the total mean score implied a normal distribution of alternation-scores.

Considering the number of alternations in each frequency series as a separate score, the number of scores less than the mean expected value exceeded the number of scores greater than the expected value for every condition. There were a total of 100 alternation-scores less than the mean expected value, 75 scores greater than the expected value, and 5 scores exactly the same as the expected value. The

[2] L. E. Moses, Non-parametric statistics for psychological research, *Psychol. Bull.*, 49, 1952, 132-133.

simplest assumption seems to be that a fairly good approximation to the expected number of alternations is given by *S*s who are instructed to construct a random series. The experiment was not, however, designed to yield statistical information regarding goodness of fit.

An analysis of variance was performed on the alternation-scores to determine the effects of symbol-frequency and sequence of presentation. The analysis of the design in Table I is presented in Table II.

Symbol-frequency, the chief comparison, was significant at the 1-% level. The sequence of presentation proved significant at the 5-% level. The mean number of alternations per condition varied from 50.30 for Sequence 1 (see Table I) to 40.93 for Sequence 3. The expected mean sequence value given by the Wald-Wolfowitz formula is 47.63. The ranking in order of decreasing number of alternations was by sequences: 1, 5, 2, 4, 6, 3. In both Sequences 1 and 5 the 70–30 condition was third in sequence.

TABLE II

ANALYSIS OF VARIANCE OF ALTERNATION SCORES

Source	*df*	mean sq.	*F*
Sequence (columns)	5	371.58	2.69*
Order-within-Sequence (rows)	2	135.14	
Symbol-Frequency	2	1065.20	7.71†
Residual (error)	170	138.10	
Total	179		

* Significant at the 5-% level.
† Significant at the 1-% level.

Summary. The supposition that mathematically unsophisticated *S*s would be unable to construct a mathematically random series, since such a constructed series would not contain enough clustering, was investigated. Sixty college students each constructed three binary random series with a length of 100 symbols apiece. The required frequencies were 50–50, 60–40, and 70–30. Six different sequences were used that order effects could be analyzed. No evidence was found to support the idea of too frequent alternation. Any slight differences found were in favor of greater-than-expected clustering. Analyzing the number of alternations for each series, it was found that the three symbol-frequencies were significantly different at the 1-% level, while sequences were significantly different at the 5-% level. With the gross level of analysis used, the prevalent *a priori* assumption that the human being is a systematically biased randomizer was not borne out.

NOTES AND DISCUSSIONS

A NOTE ON THE LEARNING OF 'SPONTANEOUS' ACTIVITY

As a measure of 'spontaneous' activity (*i.e.* uncorrelated with any specific external stimulus), the running-wheel has been used with rats chiefly to determine its genetic and physiological conditions. Early studies were concerned with such factors as diet, hormones, drugs, age, and inheritance.[1] Little attention was paid to one important condition; viz. how does the animal become a wheel-runner?

In an experiment using this apparatus for still another purpose we housed three groups of male hooded-rats in rotary activity cages and recorded the daily revolutions.[2] Unless otherwise noted, all animals were fed once a day in groups and had access to the wheels about 22 hr. a day. The results, corrected to 24 hr. and averaged in 6-day blocks, appear as Curves A, B, and C in Fig. 1.

Group A (N = 16, starting ages 91-97 days) spent every other day in activity cages from Sept. 23 to Dec. 2. For the first three weeks they were fed *ad libitum;* after the break in the curve they were put on the hunger cycle. Group B (N = 11 starting ages 75-80 days) spent 6 days out of 7 in activity cages from Dec. 4 to 29 and every day thereafter until Jan. 21. Group C (starting ages 140-143 days) lived in the cages continually. Records were complete for 8 *Ss* from Feb. 28 to April 5. Seven of them had already spent six days in the wheels on a thirst cycle; their mean score is given by the unconnected circle in Fig. 1.

As will be note, Curves A and B share two chraacteristics: a fairly steep rise followed by a slower decline. (In Group A the effect of the experience before food deprivation cannot be evaluated.) Curve C shows these properties only slightly if at all and may be considered anomalous.

To see if this type of curve was an artifact of our conditions we consulted previous studies for comparable data. Several were found in which normal male rats had served as controls for groups under experimental treatment. Three studies on the effects of castration furnished the data

---

[1] Mary Shirley, Spontaneous activity, *Psychol. Bull.*, 26, 1929, 341-365; N. L. Munn, *Handbook of Psychological Research on the Rat*, 1950, 52-83.

[2] J. P. Seward and A. C. Pereboom, Does the activity wheel measure goal striving? *J. Comp. & Physiol. Psychol.* (in press).

tween hours of food deprivation and eating. Similarly a 'drive to be active' implies a positive correlation between hours of forced inaction and subsequent activity. Such a correlation was suggested by Skinner's observations, but two other attempts failed to reveal any simple monotonic curve.[8] If there is such a drive its influence seems to be easily obscured by other factors. Perhaps 'readiness for action' may prove to be a more useful term. There is urgent need for a systematic investigation of activity as a function of previous inaction measuring both variables over a wide range of time intervals.

To explain the decline of R_w in the later portion of the curves Hull's system provides two obvious possibilities. One is conditioned inhibition (S^IR) based on fatigue reduction; the other is gradual satiation of a presumed activity-drive. A decision between the two would rest on the outcome of a test for spontaneous recovery. Over a long inactive period drive should recover; S^IR, however, should remain in force and keep activity low.

A third possibility is worth mentioning. A specific drive of wheel-running, comparable with the appetite for a specific food, may be acquired and satiated. To test this hypothesis a second group of animals 'satiated' for the wheel could be switched to some other form of activity. This procedure would prevent recovery of a general drive but would permit R_w to recover.

In conclusion, due perhaps to its history, wheel-running is often accepted as a direct expression of the organism's energy-level. The data here presented suggest that it may be fruitful to consider wheel-running not merely an index of vigor but an integrated response subject to principles of learning—not a general activity but a specific act.

University of California
Los Angeles

JOHN P. SEWARD
A. CLINTON PEREBOOM

PERCEIVED SIZE: CLOSURE VERSUS SYMBOLIC VALUE

The problem of this study was to determine whether the differences in perceived size of disks with symbols drawn upon them, reported by Bruner and Postman,[1] could better be explained by their hypothesis of symbolic

[8] B. F. Skinner, The measurement of "spontaneous activity," *J. Gen. Psychol.*, 9, 1933, 3-23; Mary Shirley, *op. cit.*, 177-184; P. S. Siegel, Activity level as a function of physically enforced inaction, *J. Psychol.*, 21, 1946, 285-291.

[1] J. S. Bruner and Leo Postman, Symbolic value as an organizing factor in perception, *J. Soc. Psychol.*, 27, 1948, 203-208.

value than by the more general Gestalt principle of closure.[2] Since figures with greater closure are phenomenally judged to be more compact and, hence, smaller than figures with less closure, we may ask whether disks bearing the dollar sign, the swastika, and the neutral symbols of Bruner and Postman are judged to be different in size from figures matched with respect to closure?

To answer this question, symbols were chosen which, though differing in form, were equivalent in closure. They had the same amount of open space and the same area in length of line. They were equated by three graduate students at Tulane University. After the symbols were selected they were drawn in black India ink on white, paper disks, 2 in. in diameter. Every disk was mounted upon a uniform, black background.

A series of comparison-disks were prepared which varied in diameter, in steps of 1/32 in., from 1 8/16 in. to 2 7/16 in. They were mounted on individual black cards. These comparison-cards were shuffled and handed to *S* who was instructed to choose from among them the comparison-disk that equalled the symbol disk in size. The distance of the symbol-disk from *S* was constant throughout the study.

Twenty-four *S*s volunteered for the experiment. Rather than use college sophomores, the services of graduate students in fields other than psychology, Tulane University maintenance personnel, and non-university individuals were obtained. Every *S* was assigned, randomly, to one of the six counterbalanced orders of presentation of the symbol-disks. Each *S* made five judgments of the size of each symbol-disk before proceeding to the next. His perception of the size of the symbol-disk was taken as the average of the five judgments.

Results may be summarized as follows. Evaluated against the uncorrelated error-term, the mean square for orders of presentation was not significant. The repeated measures, correlated components of variance, were evaluated against the pooled residual variance among the *S*s. The only source of variance which was significant was the between symbols mean square. This was significant at the 0.001-level of confidence. Having established this overall significance of differences among symbols, *t*-tests were made between symbols by pairs. There was no significant difference between pairs of symbols except when the dollar sign was one of the members of the pair. This significance was due to the size of the disk bearing the dollar sign being overestimated. The dollar sign was judged to be larger than it really was by 20 of the 24 *S*s.

[2] Kurt Koffka, *Principles of Gestalt Psychology*, 1935, 92.

We conclude that the hypothesis concerning closure was not substantiated. It should be noted, however, that the swastika produced no significant effect on perceived size, although all *Ss* recognized it. Certainly, the fact that the disk bearing the dollar sign was overestimated in size by a large majority of our *Ss* confirms Bruner and Postman's results, and indicates that symbolic value does have an effect on perceived size.

University of Illinois CHARLES M. SOLLEY
Tulane University RICHARD LEE

A NOTE ON THE MEASUREMENT OF AUTOKINETIC MOVEMENT

In a recent article, Bridges and Bitterman describe a technique for the measurement of autokinetic movement.[1] The apparatus, which is relatively simple, consists of a single lever pivoted in the horizontal and vertical planes with a point source of red light attached to one end. *S* is required to compensate for the illusory movements, *i.e.* to return the spot of light to the phenomenal center of the unstructured field. *S*'s responses are recorded on a continuous-feed kymograph, calibrated to indicate direction, extent, and presumably rate of autokinetic movement.

In the writer's opinion, this method is not an improvement over the Guilford and Dallenbach tracing method.[2] At least two sources of error are present in this technique that did not exist in the old one. They are: (a) errors caused by the interaction of real and illusory movements; and (b) tracking errors augmented by the compensatory task.

The effect of real movement on the autokinetic phenomenon has not been explicitly studied. Some investigators believe that the autokinetic movement is enhanced by the presence of real movment,[3] others that it is inhibited. To assume, however, that there are no interactvie effects between real and illusory movements is totally unwarranted. Purported measures of autokinetic movement that confound these effects are not dealing with 'pure' autokinesis.

[1] C. C. Bridges and M. E. Bitterman, The measurement of autokinetic movement, this JOURNAL, 67, 1954, 525-529.

[2] J. P. Guilford and K. M. Dallenbach, A study of the autokinetic sensation, this JOURNAL 40, 1928, 83-91.

[3] A. Graybiel and B. Clark, The autokinetic illusion and its significance in night flying, *J. Aviat. Med.,* 16, 1945, 111-151; R. H. Brown and J. E. Conklin, The lower threshold of visible movement as a function of exposure-time, this JOURNAL, 67, 1954, 104-110.

Evidence also supports the view that tracking with a compensatory display is relatively inaccurate unless target rates are slow and the course of target movement predictable.[4] Senders has shown, for example, that compensatory tracking decreases with an increase in target rates and a decrease in room illuminations. Descriptions of the autokinetic movement reveal that the pattern of movement is not highly predictable and that the rate of movement can exceed 10° of visual angle per second.[5]

The implication is clear. Techniques for measuring the autokinetic movement that do not take into account these two sources of error (*e.g.* the Bridges-Bitterman apparatus) cannot furnish a true picture of the phenomenon.

These comments suggest that the following modifications be made in the apparatus and procedure. (1) Detach the light source from the lever control. (2) In good illumination, train *S* to follow a moving light with the lever control, *i.e.* open-loop tracking. Comparisons of control and light kymograph records will show when *S* has achieved sufficient consistancy to discontinue training. (3) Present *S* with the autokinetic situation and instruct him to follow the illusory movements. Since *S* is given the same task, he should be able to track the autokinetic movement as accurately as he learned to track real movement under lighted conditions.

Montana State College JACK E. CONKLIN

ON THE INDEPENDENCE OF SUCCESSIVE RESPONSES IN THE QUANTAL PSYCHOPHYSICAL METHOD

Recent issues of this JOURNAL have seen a spirited discussion of the problem of independence of successive responses in psychophysical procedures.[1] In a recent study of the psychometric function for differential intensity sensitivity, the writer was able to gather some data pertinent to

---

[4] J. W. Senders, The influence of surround on tracking performance: I. Tracking on combined pursuit and compensatory one-dimensional tasks with and without a structured surround, Aero Med. Lab., Wright Field, *WADC Tech. Rep.* 52-229, 1953; Conklin, The Effect of Lag on Operator Efficiency in a Tracking Task, Doctoral dissertation, The Ohio State University, 1954; B. O. Hartman, An Investigation of Scale Effects in Simple, Repetitive Tracking Tasks, Doctoral dissertation, The Ohio State University, 1952.

[5] Guilford and Dallenbach, *op. cit.*, 84.

[1] V. L. Senders, Further analysis of response-sequences in the setting of a psychophysical experiment, this JOURNAL, 66, 1953, 215-228; J. E. Conklin, Senders on response-sequences, this JOURNAL, 67, 1954, 363-365.

this problem, as it specifically relates to the quantal psychophysical method.[2]

The quantal method was used to construct psychometric functions for differential intensity sensitivity at two sensation levels (10 and 40 db), at each of two frequencies (1000 and 4000 ~), in 20 *S*. Intensity increments were added to a continuous reference tone at 5.5-sec. intervals. Each increment rose from the steady-state level to maximal amplitude in 100 m.sec., remained at maximal amplitude for 550 m.sec., then decayed to the steady-state level in another 100 m.sec. *S* was instructed to press a response button each time that he was certain he had heard a change in the loudness of the reference tone.

Every stimulus-sequence consisted of the presentation of 20 increments of equal magnitude. The increment magnitude in successive sequences was varied randomly. From five to seven sequences were presented at each sensation-level of each frequency. To test for the independence of successive events within a stimulus-sequence, each event (presentation of an increment) was assigned the value 1, if the increment was followed by a response, and 0, if the increment was not followed by a response. This analysis was restricted to stimulus-sequences in which the percent response fell within the range from 20 to 80%.

One measure of the randomness of such a sequence of binary digits is the serial correlation coefficient with lag one.[3] This is a non-parametric statistic based on the cross product term

$$R = \sum_{i=1}^{n} x_i x_{i+1}$$

where X_i is the outcome of the *i*th event in the sequence, and *n* is the number of events in the sequence. The value of *R*, under the null hypothesis, varies as a function of both *n* and the percent response, but, for large *n*, it is approximately normally distributed with mean and variance which are functions of *n* and the percent response. Thus, values of *R* obtained from all sequences may be compared by the transformation.

$$Z = R - E.(R)/\sigma_R.$$

An analysis of variance failed to demonstrate that either frequency or sensation-level had a significant effect on *z*. Hence, all *z* values for a single *S* were averaged. This yielded a distribution of 20 mean *z* scores. Under the null hypothesis (no serial correlation) one would expect this distribu-

[2] S. S. Stevens, C. T. Morgan, and John Volkmann, Theory of the neural quantum in the discrimination of loudness and pitch, this JOURNAL, 54, 1941, 315-335.

[3] P. G. Hoel, *Introduction to Mathematical Statistics*, 1947, 182.

tion to have zero mean and unit variance. The observed sample mean was +0.056, the standard deviation, 0.65. The deviation of the sample mean from zero is not sufficient to reject the null hypothesis ($t = 0.376$).

These results seem to suggest no lack of independence of successive responses in the quantal psychophysical method, at least for sequences of 20 increments. It should be noted, however, that, in the present study, *S* was given no visual warning signal prior to the presentation of each increment. Since visual warning signals have been used in some studies employing the quantal method, it would be interesting to know whether independence is preserved in this latter situation.

Northwestern University JAMES F. JERGER

THE 1954 ANNUAL MEETING OF SECTION I, AAAS

The one hundred twenty-first meeting of the American Association for the Advancement of Science was held in Berkeley, California, December 26 through December 31. Section I (Psychology) arranged six sessions, three of invited papers and three of submitted papers. Section I also co-sponsored six additional sessions which were arranged by the Western Psychological Association, The Society for Research in Child Development, and the Third Berkeley Symposium on Mathematical Statistics and Probability.

John P. Seward organized a symposium to deal with the subject "How can behavior theory best handle the construct of motivation." T. W. Forbes arranged a session of invited papers on the topic, "Psychological factors in highway safety." This session was co-sponsored by Section M (Engineering) and discussions of the formal papers presented were led by representatives of the engineering and legal professions. The third group of invited papers, arranged by Donald B. Lindsley, was concerned with problems of "The nervous system and behavior."

The Society for Research in Child Development planned and presented two groups of papers. The topics considered were: "The values and limitations of longitudinal studies of children" and "Approaches to the study of personality development in children."

The Western Psychological Association arranged three symposia, the subjects of which were: "The present status of psychoanalytic theory"; "Sex differences in personality and intellectual development"; and "Perception: Learned and unlearned."

The Vice Presidential Address of Section I was given by Donald B. Lindsley on "The crossroads of psychology and neurophysiology," after summarizing the results of recent electrophysiological studies of the reticu-

lar formation and diffuse thalamo-cortical projection system, he discussed the application of these results to psychological problems.

Section I joined with Section K (Social and Economic Sciences) in sponsoring a social sciences dinner meeting at which John B. Condliffe, Professor of Economics, University of California spoke on "The international consequences of scientific research."

University of Chicago W. D. NEFF

A LIMITATION OF THE 'MAXWELLIAN VIEW'

The very useful note by Dr. Leibowitz on "The use and calibration of the 'Maxwellian View' in visual instrumentation," which appeared in the September 1954 number of this JOURNAL,[1] failed to mention a very important limitation to the use of the Maxwellian view, occasioned by the Stiles-Crawford effect.[2] These authors found that the luminosity produced by a narrow beam entering the eye and striking the retina depends on the region of the pupil through which the beam passes, being greatest when it enters at some point near the center of the pupil. The effect is not negligible, since a displacement of 2 mm. may vary the luminosity by a factor of two. This also makes the focusing of the beam on the eye a criitcal matter, since, as Collier has shown,[3] an image 2 mm. in diameter may produce a luminosity 10% less than one 0.5 mm. in diameter or smaller. Very careful focusing and accurate positioning of eye are therefore essential in the use of the Maxwellian view.

Ottawa, Canada W. E. KNOWLES MIDDLETON

TABLES FOR COMPUTING INFORMATIONAL MEASURES

The Operational Applications Laboratory of the Air Research and Development Command has published a set of tables for computing informational measures. This includes: (1) $\log_2$ n for whole numbers from one to 1000; (2) n $\log_2$ n for whole numbers from one to 500; (3) — p $\log_2$ p for four place probabilities between 0.0001 and 0.2500 and three place probabilities between 0.251 and 1.000; (4) formulas for computing information measures; and (5) an illustrative example. A limited number of these tables are available for free distribution. Address requests to the Operational Applications Laboratory, Bolling Air Force Base, Washington 25, D.C. K. M. D.

---

[1] Herschel Leibowitz, *op. cit.*, this JOURNAL, 67, 1954, 530-532.

[2] W. S. Stiles and B. H. Crawford, The luminous efficiency of rays entering the eye pupil at different points, *Proc. Royal Soc. London* B112, 1933, 428-450.

[3] L. J. Collier, Visual telephotometry, *Trans. Illum. Eng. Soc.* (London), 3, 1938, 141-154.

Harvey A. Carr: 1873-1954

Harvey A. Carr was born on a farm in Indiana on April 30, 1873, and died at Culver, Indiana, at the age of eighty-one years, on June 21, 1954. The middle initial in his name had no significance. It was adopted and used by him for the purpose apparently of rounding out his signature. He had his first training in the country schools of Indiana. At eighteen years of age he entered the preparatory department of DePauw University and continued through two years of the college course. Then ill health and financial pressure sent him back to the farm. When twenty-six years old he enrolled in the University of Colorado, where he became interested in psychology in the classes of Professor Arthur Allin. After receiving the A.M. degree in 1902 he was given a fellowship at Chicago where he worked with Dewey, Angell, and Watson, attaining his doctorate in 1905. As he did not immediately secure an appointment in psychology, he taught for a year in a high-school in Texas, and then for the two succeeding years he was an instructor at Pratt Institute. After that he was called back to Chicago as an assistant professor when in 1908 Watson left for Johns Hopkins. He continued at Chicago in different ranks until his retirement in 1938. He was chairman of the department for sixteen years from 1922 to 1938.

Carr's publications covered a wide range of subjects. His doctoral dissertation, "A visual illusion of motion during eye closure," started one line of investigation running to the autokinetic illusion and others.[1] It culminated in his work on space perception which was published in 1935.[2] The thesis foreshadows the care in experimentation and caution in interpretation that characterized all of his work. The book has a slightly different approach to the subject from that of Helmholtz, Hering, and the Gestalt psychologists, although it covers much the same field. One critic remarked that the only reference in it to Hering is use of one illusion. Carr was little interested in the problem of nativism and empiricism which he regarded as insoluble. He restricted himself mostly to statements of phenomena and their conditions.

In 1908 Carr took charge of the work in animal psychology that Watson had started. He devised a maze that has been widely used and he made several studies of the senses that the white rat uses in running the maze, comparing normal animals with those deprived of a sense. Sight apparently contributed most, but other senses aided. Carr's presidential address before

---

[1] *Op. cit.*, *Psychol. Monog.*, 7, 1906 (No. 31), 1-127.
[2] *An Introduction to Space Perception*, 1935, 1-413.

the American Psychological Association in 1927, "Interpretation of the animal mind," discussed at length the evidence for the assumption of consciousness in animals.[3] He concluded that the only basis for a decision lies in the similarity between the responses of men and of animals, and was doubtful whether the similarity is sufficient to justify a conclusion. In this address he asserted, "I am somewhat of a behaviorist in the field of animal psychology, although I do not class myself as such so far as human psychology is concerned."[4] This he said in spite of his expressions of personal approval of Watson.

The same caution on other problems is shown in several papers published in the 1930's. In one he questioned whether the results of experiments indicate that there are absolute realities or values that are concealed by varying conditions or errors or whether results are averages convenient in stating results for a few individuals. Especially did he question the range of attention, limens, curves of forgetting, plateaus.[5] In another article he queried the existence of emotion, memory, attention, and much of the standard vocabulary of chapter heads. He complained that the words have been taken from the popular vocabulary and come to imply the existence of reals. This is a renewal of Carus' problem of reification. Carr made, however, no attempt to supply substitutes that would permit psychologists to express themselves without false implications.

Much of Carr's influence upon psychology came through the work of his students in their dissertations at Chicago and later when they were themselves directing research at other universities. In his autobiography Carr asserts that during the connection with the Chicago department 131 degrees were given to students who came under his influence in some way. He was directly responsible for 18 doctoral dissertations in animal work, 6 in space perception, and 29 in learning. Chicago was one of the most productive universities in psychology.

In his autobiography Carr also raises the question as to whether psychology exists as a science. He was doubtful at times as to whether experiment is the only method in psychology, but was optimistic, on the whole, about the progress of the science and psychology's claims to be a science. In this connection he asked what he has himself contributed to the science, but modestly refrained from giving an answer. Now that his task is finished one might raise the question again, even if an answer can-

---

[3] *Op. cit., Psychol. Rev.,* 34, 1927, 87-106.
[4] *Ibid.,* 104.
[5] The quest for constants, *Psychol. Rev.,* 40, 1933, 514-532.

not be quite definitive. There can be no question that Carr contributed to psychology a large number of important facts and developed some useful methods. His summaries in his two books are useful, and his cautious attitude is invaluable. If we consider also the work of his students during their training and later, the answer to his own question must be that his contributions were important.

Advanced students who came under Carr's influence were impressed by his kindliness, his freedom from dogmatism and his fundamental honesty, as well as by his comprehensive knowledge and his ingenuity in experimenting. All of them appreciated his caution. Most of them profited from his constant concern that some conclusions must always be compatible with experimental results.

University of Michigan

W. B. PILLSBURY

AN ACKNOWLEDGMENT

The photograph of Harvey A. Carr reproduced in the frontispiece of this number was obtained from his wife, Mrs. Antoinetta Carr. It was taken when Dr. Carr was sixty-five years old, shortly before his retirement from the faculty of the University of Chicago in 1938. The signature, a facsimile of which appears under the portrait, was also supplied by Mrs. Carr. It is from a legal document signed by him in 1952. It is a rare sample, as he usually abbreviated his given name or if he wrote it in full he omitted his middle initial. The JOURNAL's thanks are due Mrs. Carr for her helpful coöperation.

K. M. D.

BOOK REVIEWS

Edited by M. E. BITTERMAN, University of California

Modern Learning Theory: A Critical Analysis of Five Examples. By WILLIAM K. ESTES, SIGMUND KOCH, KENNETH MACCORQUODALE, PAUL E. MEEHL, CONRAD G. MUELLER, JR., WILLIAM N. SCHOENFELD, and WILLIAM S. VERPLANCK. New York, Appleton-Century-Crofts, 1954. Pp. xv, 379.

This book is the product of a group of seven participants in the Dartmouth Seminar on Learning Theory. The group, sponsored by the Social Science Research Council, took as its task the critical evaluation of the learning theories of five promient psychologists: C. L. Hull, E. C. Tolman, B. F. Skinner, Kurt Lewin and E. R. Guthrie. The specific criteria for evaluating the theories were chosen in the belief that "wide agreement would obtain among current writers in the logic of science that an adequate review of any scientific theory must include essentially the same features" stressed in this book. Consequently, the discussion of the theories centers in large part about their formal properties. The authors are aware that theories can be evaluated against any number of criteria, such as the historical context in which they appeared, the extent to which they have stimulated research, and the like. These considerations, however, play a minor rôle in the book.

The authors' justification for the book is the belief that "the strictest criticism of current formulations is necessary to facilitate further development," and that psychologists "may expect to learn as much from the errors and the failures of these attempts [at theory construction] as from their successes." Seemingly the implicit assumption the authors make is that further advances in psychology will occur only if psychologists become increasingly aware of the methodological requirements for theory construction and show far more willingness than at present to meet these requirements in formulating their own theories and hypotheses.

To paraphrase the evaluative criteria used, the authors believe that all attempts to generalize about empirical data, whether at the level of hypotheses or theory (no explicit difference between these terms being made in the book), involves at least the following steps. A set of concepts, or constructs, and their interrelationships must be stated in such a way that logical implications can be deduced from them, and it must be possible to test these implications experimentally to determine their truth or falsity. The range and clarity of the implications that can be derived must vary with the independence and consistency of the underlying concepts, the quantitative level at which the interrelationships among concepts are stated, and the like. These derived implications, moreover, must be stated in terms of the systematic dependent and independent variables of the system, *i.e.* the variables which the system is designed to relate, such as stimuli and responses. An unambiguous test of these relationships is only possible, however, if the systematic variables can be univocally coördinated with the operations and observations of the experiment. Thus, the testability of a theory depends in large part on how the concepts are related

to the systematic variables and on how these variables are related to the data-language used in describing the experiment. A full evaluation of a theory demands consideration of its adequacy in relation to each of these points.

The allotment of space to the five theories varies greatly. Koch devotes 176 pages to Hull, MacCorquodale and Meehl review Tolman in 90, Verplanck summarizes Skinner in 50, Estes evaluates Lewin in 28, and Mueller and Schoenfeld dispose of Guthrie in 34. It is not clear whether this allotment is dictated by the idiosyncrasies of the authors or whether it is a group decision concerning the relative importance of the various theories. A brief sampling of the content of each chapter follows.

(*1*) Koch, in his review of Hull's *Principles of Behavior* and *Essentials of Behavior,* concludes that (*a*) none of the systematic independent and dependent variables in the theory have an unambiguous relationship with the intended range of reductive symptoms; (*b*) no given intervening variable is securely and univocally related to its relevant systematic independent or dependent variables; and (*c*) no given intervening variable is related to any other intervening variable with sufficient determinancy to permit quantitative passage from one to the other. He documents these conclusions with extended detail. In discussing the *Essentials,* he advances the interesting and plausible hypothesis that the extensive changes in theory it contains were dictated solely by Hull's attempt to quantify reaction potential, rather than by any sensitiveness on Hull's part to the implications of experimental data, such as the studies of latent learning. While this view might account for the changes in interrelationships among constructs, it seems unlikely that it accounts for the inclusion in the system of a number of new intervening variables, *e.g.* stimulus-intensity dynamism, incentive motivation, and the like.

(*2*) The chapter on Tolman, by MacCorquodale and Meehl, contains a heterogeneous set of material. While the first 22 pages deal with Tolman, the remaining 60 pages of text contain an extended discussion of the latent learning experiments, a digression on the definition of a response, and a limited and preliminary set of postulates designed to formalize expectancy theory. The brief analysis of Tolman is attributed to the extremely programmatic nature of his theory and its lack of formal content, which permit only a restricted application of the methodological criteria in which the authors are interested. In their preliminary attempt to formalize expectancy theory, they define *expectancy* in terms of an elicitor, a reaction-class, and an expectandum. It is symbolized as *SRS* in keeping with Hull's ${}_sH_r$. The two *S*-terms correspond to the usual usage, except that the second always has some degree of valence. Valence is defined as a function of the organism's "need" and the cathexis of the stimulus, *i.e.* the number of contiguities between it and a consummatory response. The postulates concerning the acquisition, extinction, generalization, and activation of these expectancies, bear, on a casual reading, a family resemblance to Hull's postulates concerning habit-strength. These postulates, however, permit generalization between different elicitors, between elicitors and expectanda, and between different expectanda, with a resulting transfer of valence. There is thus the suggestion that they may mediate experimental results favorable to Tolman's point of view.

(*3*) In reviewing Skinner's *The Behavior of Organisms,* Verplanck finds that this system "rather than being a set of empirical laws embodying statements that

represent inductive generalizations" is actually "a set of formally defined terms, and defining laws, which are only coördinated with data-language statements after they have been fully stated." He believes that Skinner's formulations and methods are determined in large part by a set of implicit axioms, one of the most important of which is the assumption that the true laws of behavior always are characterized, even for the individual organism, by smooth and orderly functions or graphs relating the dependent and independent variables. Verplanck believes that this axiom accounts for Skinner's definitions of *stimulus, response,* and *reflex,* his use of constructs such as the *envelope of an extinction curve,* his tendency on occasion to offer *ad hoc* explanations of why some samples of behavior fail to show this orderliness, and the restricted range of behaviors and experimental variables with which he has dealt. The impression is given, however, that Skinner's formulations more nearly have met the methodological requirements of theory building than have those of his contemporaries.

(*4*) At first glance it appears strange that, in a book devoted to learning theory, Estes should have chosen Lewin as most representative of the field theorists, since Lewin's only direct discussion of learning is in his very early work and his brief chapter in the *NSSE* yearbook. Estes justifies this choice on the ground that among field theorists only Lewin's system of behavior is stated explicitly enough and in sufficient detail to merit a methodological evaluation. The main weakness that he finds in this system is Lewin's tendency to stop at "psychological" concepts with little or no coördination of these concepts with observable events or behaviors. As a result, the system can be used only to represent or describe situations and has little predictive power. At the same time Estes recognizes Lewin's contribution in showing the necessity for, and the value of, elaborating the properties and interrelationships of the concepts within a theory.

(*5*) Mueller and Schoenfeld believe that Guthrie has more a program for a theory than a formal theory in any real sense. They stress that the simplicity of Guthrie's principles and explanations is more apparent than real, a fact which they attribute to his ambiguous and conflicting usages of the terms *stimulus* and *response,* and to the numerous hidden assumptions underlying his statement of principles. The longevity of this system, despite its failure to develop with advances in empirical knowledge, they account for in part by Guthrie's insistence on a monistic view of learning, a thesis that appeals to many psychologists, and in part by the strength derived by the theory from its overlap with other theories of learning.

As an evaluation of these theories in terms of their formal and methodological properties, this book is unique. It should establish a standard for such reviews in terms of its relative thoroughness, dispassionate tone, and tendency to stress important problems, such as the universal difficulties psychologists have in defining and using such terms as *stimulus* and *response* in a consistent manner, in attempting to state truly quantitative laws, and in seeing clearly the actual assumptions underlying their own hypotheses. It provides a fairly extensive outline and illustration of the types of question that need to be asked in such reviews, and, by demonstrating the limitations and ambiguities of each of the theories, it possibly may have the therapeutic effect of decreasing ego-involvement in them to such an extent that a clearer discrimination can be made between meaningful theoretical controversies and those arising out of ambiguities in the original formulations. For these reasons,

the book should be of considerable value to workers and advanced students in the field of learning. If it is to be read with profit, however, a familiarity with the original sources is necessary, for there is little attempt at exposition of the theories evaluated. It places a considerable strain on one's apperceptive mass.

There is a sense, however, in which the conclusions reached in this book are trivial. No one familiar with the field will be surprised to learn that a psychologist racked against these criteria comes out bruised and battered. Even the specific criticisms have a familiar ring because of the extensiveness with which the field has been reviewed in the past. The present reviewer's feeling of disappointment stems from the general negativeness of the result. Just how this book advances the authors' intent "to facilitate further development" in theory construction is never made clear. The positive guideposts have not been underscored. The reader is given the impression in one chapter, for example, that psychologists should forgo the folly of trying to formulate general theories of behavior and aspire no higher than restricted hypotheses or miniature systems. In the very next chapter there is a new attempt to formulate a general expectancy theory. Thus it goes.

Perhaps this concern over the negativeness of the conclusion is an expression of the perennial conflict between the tough-minded and the tender-hearted. More likely, however, it represents a difference in opinion concerning the rôle of theory in the extension of knowledge, and the relative value of the various functions a theory can serve. The conflict seems to center about the contrast between the 'fruitfulness' of a theory and the logical rigor with which it is stated. The authors at times seem to feel this same opposition. Koch, for example, frequently refers to the ingenious and fruitful "pre-theoretical" hypotheses advanced by Hull. There can be no doubt, however, that, if the present criteria concerning formal rigor were applied to them, they would disintegrate. Again, Koch condemns the increased "localism of the theory language" he finds in Hull's *Essentials of Behavior* because it restricts the generality of the formulation; but a plausible argument could be made that this self-defeating specificity resulted from an overemphasis on rigorousness. In any event, the book leaves the impression that, if there really is on occasion a conflict between fruitfulness and formal rigor, priority should be given to the latter—that only ideas stated fully and unambiguously can advance research in the long run.

There can be no quarrel about the desirability of logical clarity and rigor in the formulation of ideas, but at best these logical demands are only a necessary and not a sufficient condition for a worthwhile theory. In present day psychology many of the most rigorous ideas seem most empty of content, and it is the nature of the content that determines the degree of fruitfulness. Perhaps the fact that the latter can only be judged historically, and even then is hard to evaluate, explains why the scientific methodologist places so much emphasis on rigor. It is so much easier to judge.

An alternative way of stating the case is that the book seems to hold up an ideal limit for theorizing. In a sophisticated manner it suggests that the ideal situation is one in which each experiment gives a *yes* or *no* answer to a strict logical deduction from the theory. Like the purple cow, such a situation has never been encountered, and probably no one would want to encounter one even if he could. This ideal situation would imply, first of all, that the field under consideration had been

so exploited that in large part the theory now served only the function of describing known relationships, and, secondly, that the theorist was completely aware of all the assumptions underlying his hypotheses. Only then would the theorist be in position to make completely rigorous, unambiguous deductions that had any bearing on empirical fact. As a consequence, experiments designed to test the hypothesis would carry little information; they could only answer *yes* or *no*. This would be the point at which everyone lost interest in the proceedings, with the possible exception of the logician lost in esthetic contemplation of the result.

In practice, the value of an experiment lies as much, if not more, in what it suggests than in what it proves. The reason an experiment contains extensive information is to be found in the uncertainty that gave rise to it. This uncertainty is not merely a question of two alternatives, yes or no, true or false. The uncertainty is much broader, and it stems from the ambiguity of the concepts underlying the hypothesis being tested. Each concept always has in some degree a penumbra of ambiguity as to what it really means. Concepts that eventually prove to be fruitful always have an initial ambiguity, although the converse need not be true. Their fruitfulness in part consists of the fact that they are open to a wide range of alternative interpretations, and thus may be aplicable to a wide range of problems. One function of the experiment is to help narrow this range, but that is not its only function. No one can exhaustively list these alternatives prior to experimentation. The theorist never knows what he really means by his own ideas until he sees them in action. Paradoxical as it sounds, experimentation serves simultaneously to enlarge our awareness of the meanings of a concept while narrowing the range that is actually applicable in a given case. The former is what is meant by the suggestiveness of an experiment.

Of necessity then, there must always be inconsistencies, ambiguities, and logical gaps in any theory that is relevant to an ongoing field of research. The function of research is as much that of clarifying this ambiguity of our concepts gradually, continuously, and empirically, as it is the testing of well formulated hypotheses. Logical fiat and strictures about clarity are no substitute for this approximating procedure. If there is any value in this attitude, it enjoins us to develop a tolerance for such ambiguity and to demand no more rigor of a theorist than is commensurate with protecting the true intent of his concepts. Premature logical rigor in a system may all too frequently prune away the alternatives available to a concept before they have even been discovered.

These remarks are, however, tangential to the main core of the book. The authors took as their task the evaluation of present theories in terms of a set of criteria proposed by writers on the logic of science. This task they have accomplished admirably. It is only if they have taken seriously their secondary purpose, of facilitating further advances in theorizing, that one can quibble with their approach to the problem. Good models for behavior are only copied when they can be discriminated and one is reinforced for imitating them. Fechner's principle suggests that advances in psychological theorizing would be more discriminable when judged against previous practice than when judged against the endlessly distant ideal limit. Similarly, learning theory itself suggests that more is acquired from positive instances than from negative ones.

Stanford University

DOUGLAS H. LAWRENCE

Biology and Language. By J. H. WOODGER. Cambridge, The University Press, 1952. Pp. xiii, 364.

Professor Woodger delivered the Tarner Lectures in 1949-50 at Trinity College, Cambridge. His seven lectures, supplemented by notes and appendices, comprise this book.

Woodger is probably best known to American psychologists for the influence he had on the work of Clark Hull and on the use of the axiomatic method in the mathematico-deductive theory of rote learning. Woodger's two previous books apply the axiomatic method to biology, particularly to genetics, but he presented that work in large measure as a specific illustration of how the method should be applied to the general task of theory construction. Whether or not the Yale group profited from Woodger's stimulation is a debatable question, but it is interesting that the attempt to rephrase every statement in symbolic logic has not been tried again. One gathers that the logical formulation involves more effort than profit. Many psychologists have come away from Woodger and Hull with the impression that the price one must pay for a rigorous symbolic theory is to make the obvious obscure and the obscure impossible.

Apparently Woodger has not been insensitive to this reaction, for in the present book he has greatly simplified and clarified his procedures and has dealt much more explicitly with the way theoretical statements are based on observation and experiment. The result is, in most parts, a more readable discussion that will appeal to a much wider audience. Readers too lazy to decipher a formula will still be excluded, but the whole conception of this book is larger, clearer, and more persuasive than any of Woodger's previous writing.

The lectures are addressed to biologists, yet they are not about biology. More precisely, they are not about organisms. The lectures are about the statements that biologists make about organisms and the relations of such statements to one another and to empirical observations. Woodger holds a low opinion of the native linguistic habits of most biologists and in these lectures he tries to show how muddles have resulted from vague and ambiguous statements, and how such muddles can be avoided by the biologist who is familiar with scientific methodology and the axiomatic method.

The book is divided into three parts. Part I, on general methodology, contains two lectures and 92 pages. Part II, on methodological problems in genetics, has three lectures and 157 pages. Part III, on methodological problems in neurology and related sciences, has two lectures and 101 pages. It is the third section that will probably be most interesting to psychologists, for here Woodger comes to grips with some psychiatric problems and has some suggestions to make about a method of formulating statements for a theory of behavior. The introductory material in Part I has been better introduced elsewhere and the beginner will not find this the best place to start his logical education. It is, however, an interesting statement of Woodger's philosophical position, nominalism. The application to genetics in Part II provides the best illustration of what Woodger is recommending, but it is quite technical and not of direct interest to psychologists. Consequently, the following appraisal will be confined to Part III. It is almost but not quite possible for the reader who knows a little formal logic to understand all of Part III without looking at the preceding 250 pages.

Woodger observes that neurologists and students of animal behavior fall into two classes: those who hold that the language of neurology and behavior should be strictly physicalistic, and those who freely mix non-physicalistic words into their scientific language. Questions concerning the legitimacy of such a mixed language traditionally lead to a discussion of the body-mind dualism, and this discussion is usually aproached from the point of view of the alleged entities that 'body' and 'mind' are supposed to name. The question of how these words function in various scientific uses is seldom considered. Woodger proposes to take a linguistic or methodological approach—to analyze the statements we make in order to discover different classes of words and the rules for using them.

For example, consider the statement, "John got a view of the sea," and ask what words might be substituted for *John, view,* and *sea.* That is to say, consider the linguistic frame, "*x* gets *y* of *z*." The values of *x* are names of persons *John, you, Mary, they, people,* the values of *y* are names of sensible objects (*view, sound, feel, smell, taste*), and the values of *z* are either names of persons or names of physical objects (*sea, tree, hours, vinegar, airplane*). It is clearly impossible to substitute a word from one class for a word from another class. One would not say that "A view got John of the sea" or "The sea got a view of John" or "John got a sea of the view." Thus there are here three mutually exclusive classes of words. The names of sensible objects (which is Woodger's neutral term for sensations, or perceptions, or cognitions) stand linguistically in one relation to the names of persons and in another relation to the names of physical objects. In this linguistic or methodological observation the question of body-mind dualism need not arise. Now when a neurologist says that the eighth cranial nerve carries sensations of sound to the cortex, we do not criticize him for mixing the 'mental' sensation with the 'physical' nerve and cortex; we criticize him for his violation of the linguistic rules. He cannot say that a brain gets sensations of sound any more than we can say that the sea gets a view of John. We must say that sensations of sound are gotten by persons, not by ears or nerves or brains.

A theory that uses only the person language is said to be animistic; a theory that uses only the sensible-object language is said to be phenomenalistic; and a theory that uses only the physical-object language is said to be physicalistic. Behaviorism is a physicalistic theory because it holds that the study of behavior should be confined exclusively to the physical-object language of observable, overt behavior. The behaviorist replaces a person by his body and replaces his sensible objects by his bodily movements. Woodger takes issue with those who argue that the behavioristic theory is the only possible one for scientific psychology. It has not been proved that a behavioristic psychology will be successful. There is no insurance that many important psychological issues can survive such physicalistic substitutions. Nor has it been shown that a psychological theory in person language is impossible.

Instead of the traditional dualism of mind and body, Woodger proposes that we deal with four different languages: physical language, the language of sensible objects, person language, and community language. In ordinary English all these languages are hopelessly muddled together. The physical language is the most easily sorted out, but with the others there is chaos. Rules for their construction and use in scientific theories are almost completely lacking. Given such axiomatized theories it might be possible to discuss intelligently whether a theory in, say, person lan-

guage was 'reducible' to a theory in physical language, but at present this question is quite meaningless.

Woodger presents some of the ideas of I. D. Suttie to illustrate what a theory in person language would look like. To the reader excited by the prospect of extending the Boole-Frege techniques into such obscure and difficult areas, the result is a melancholy sight. It is not that Suttie's ideas are wrong or stupid, but rather the discouragement lies in the fact that their present form is so verbal, intuitive, and slippery that an axiomatic formulation looks exceedingly remote.

The seventh lecture is not up to the rest of the book. The sixth, which presents the idea of four languages, is well worth studying, but one gets the impression that Woodger was running out of material when he got to his last lecture, and that he fell back on some rather unrelated opinions about bad manners in the medical profession, some reminiscences about his old friend Suttie, and a rather uncritical division of mankind into the tough-minded, physicalistic, behavioristic, totalitarian, egocentric propagandists versus the tender-minded, gentle, humanistic, democratic, sociocentric citizens. These opinions are certainly interesting enough, but they are not always persuasive. The conclusion that applied behaviorism must lead to a totalitarian state is especially superficial. Apparently the caution Woodger shows in genetic questions is not a general trait.

This is a spotty book, but if you are willing to study the right spots you are reasonably sure to get some original ideas about some old problems.

Massachusetts Institute of Technology GEORGE A. MILLER

The Science of Color. By THE COMMITTEE ON COLORIMETRY, OPTICAL SOCIETY OF AMERICA. New York, Thomas Y. Crowell Company, 1953. Pp. xiii, 385.

This book is the final report of the second Committee on Colorimetry of the Optical Society of America. The task of the committee was to revise and bring up-to-date the first report which appeared in 1922 under the chairmanship of the late L. T. Troland. Needless to say, there have been many advances in the years intervening between the two reports. As it stands, *Science of Color* is more than an ordinary report on colorimetry. It attempts to cover almost every aspect of color—not merely matters relating to terminology, methods of measurement, and problems of standardization. The range of topics is very large indeed: from the art of coloring in ancient civilizations, and the anatomy and physiology of the eye, to psychological, physical, and psychophysical concepts of color as well as very technical problems of colorimetry and color-standards. From the point of view of completeness and, in many respects, in detailed treatment, this book is a significant contribution to the vast literature on visual science. The committee responsible for the contents consisted of 23 specialists from the various fields of color of whom six wrote the original chapters or parts, and two, L. A. Jones, chairman of the committee, and D. L. MacAdam, revised and were ultimately responsible for what was actually published. S. M. Newhall wrote the original material on psychological aspects of color. Since other members of the committee contributed criticisms and suggestions, the book is a coöperative work which has the excellences and defects usually found in symposium volumes.

Although part of the report appeared in the Journal of the Optical Society in 1937, it was not until 16 years later that full and final publication was achieved in

the present volume. The reason for the long delay was "a stalemate between the physical and psychological viewpoints . . . within the committee." It was not until differences concerning fundamental definitions of light and color were resolved that the full report could be published. In the first report, color and light were regarded as sensations, or purely psychological, whereas some members of the committee felt that they should be purely physical. The compromise finally reached was to make them both physical and psychological, or psychophysical; but psychophysics as defined here does not mean what it means to psychologists, for psychophysical measurement is restricted to "specifications of stimuli in accordance with equality or difference of sensations" (p. 59). According to this point of view, many psychophysical methods currently in use by psychologists are not truely such, *e.g.* the method of limits, the method of paired comparisons, the ratio-method, fractionation methods, and the methods of absolute judgment and of comparative ratings. Yet the narrow view does not entirely prevail, for we find Newhall discussing methods and results that do not fall under the definition of psychophysics. In fact, the long struggle over fundamental definitions seems to have had little influence on choice of material, since the point of view underlying the definitions is so restrictive that had it prevailed much of the most valuable material on color would not be in this book.

According to the strict definition, color is limited to the non-temporal, non-spatial aspects of vision and has only three attributes—hue, brightness, and saturation. Yet it is admitted that extent and duration do affect the three attributes (p. 105). If the definition had been strictly adhered to there should be nothing in the book about contrast and adaptation effects, or about flicker. On second thought, there should be nothing about color in the book, because even points of color are seen against extended surrounds!

The psychological aspects of color are divided into sensation and perception. The reason given for this division is a statement from Katz to the effect that "different modes of appearance of the same color are all based on the same retinal processes" and hence the attributes of hue, brightness, and saturation may be regarded as an irreducible, constant residuum that is color in the strict sense. Whether or not Katz meant what some members of the committee here thought he implied is an open question, but it can be stated categorically that so-called modes of color-appearance do affect hue, brightness, and saturation, and no one can assert that the retinal processes are the same when the attributes are different from any cause whatsoever. The film or aperture colors, which some physicalists and, one is sorry to say, even some psychologists, regard as atomistic bits of color unchanged by the context in which they appear, are pure abstractions. All the evidence points to the contrary. Aperture colors are as much subject to surround-effects as surface colors. An aperture color can be made brighter or darker, of one hue or another, and of any saturation by changing the surround as the reviewer has shown in numerous publications. One can specify a color as having a given hue, brightness, and saturation only under one specific set of conditions, and no one set of conditions has priority over another so far as basic color-quality is concerned. That a broader basis for colorimetry is necessary for the purposes of color-specification in industry and science than is afforded by the C.I.E. standard observer must be patent to all. The standard observer should be defined not only for small foveal stimuli against

completely dark surrounds but also for neutral light surrounds, chromatic surrounds at photopic levels, and for completely scotopic conditions of viewing.

It is evident that the committee was plagued by philosophical considerations that most workers in visual science have learned to side-step in order to get ahead with their scientific problems. If it is possible to measure phenomena of color accurately, what difference does it make whether they are called sensations or perceptions? Certainly one distinction should be made in color work, as in all other fields where organisms are concerned—the distinction between stimulus and response. Often response is taken for stimulus as in this book when it is stated that "the retina must respond in at least three different ways to *different colors*" (p. 91) when what was meant was *different spectral energies*. The fact that responses differ in complexity does not necessarily put them in different universes of discourse. It is to be hoped that the next report on colorimetry will adopt definitions more in keeping with the actual status of scientific accomplishment.

Although the book as a whole is authoritative and sound, there are many statements in it which call for revision. Among the more important ones to this reviewer are the following:

In the discussion of the modes of color-appearance in the Introduction, reference is made to Troland's treatment published in 1929 after which it is stated: "David Katz in his book published in 1935 . . . treats the same modes of appearance as those mentioned by Troland" (p. 7). This discussion of the modes implies that Troland preceded Katz. On page 145, it is correctly stated that Katz was the first to describe the modes in a systematic way. While nothing can be said to add to or detract from the glory of both these great workers in the field of visual science, this correction should be made at least for the sake of historical accuracy.

The statement that "It is obvious that memory plays an important part in the assignment of color to an object" (p. 47) is too sweeping. It is interesting that the title of the first edition of Katz's classic work was *Die Ercheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung*, but in the second edition the subtitle referring to individual experience was omitted because Katz admitted he was unable to find any experimental evidence for the influence of memory on color. Memory may play a part in some experience of color under certain conditions, but it has too often been assumed to be responsible for color-effects when it was not responsible for them.

The reviewer cannot agree with the statement that "the weight of experimental evidence tends to reaffirm the binocular production of yellow" (p. 123). One of the references quoted in support of this statement, a paper by Jameson and Hurvich, presents cogent experimental evidence against binocular yellow in Hecht's sense; and the second of the three references cited in support of the statement (Prentice) states explicitly that even more careful control of stimulus-conditions than he was able to exercise will be necessary before it can be said that yellow is a product of brain action alone.

In discussing the effects of colored backgrounds on the colors of stimuli, the committee states that "For the most critical work even very slightly chromatic backgrounds should be avoided" (p. 51). This statement may be true for some types of critical work, but it is certainly not true for all types of color work. D. C. White has shown that, if it is desired to maximize the difference in color

between stimuli of not too dissimilar hues, a background having a hue intermediate to that of the stimuli will increase the color difference (Stanford University Dissertation, 1949).

All in all, this book is the most comprehensive treatment presently available in the field of color. The 25 colored plates, 102 figures, and 40 tables contain a wealth of illustrative and quantitative data that cannot be duplicated between the covers of any other book. It can save much labor in providing ready access to data that must be sought in various periodicals, separately published manuals, and numerous odd sources. It seems supererogatory to ask for more in a book that already contains so much, but this reviewer misses data on and discussion of fluorescent sources which are being used more and more for all sorts of purposes. To those who, like the reviewer, offer a course in vision and want a text that is modern and fairly representative of the colorimetric approach, this book will surely appeal.

University of Texas HARRY HELSON

Studies in Schizophrenia: A Multidisciplinary Approach to Mind-Brain Relationships. By the TULANE DEPARTMENT OF PSYCHIATRY AND NEUROLOGY. Reported by ROBERT G. HEATH, Chairman. Cambridge, Harvard University Press, 1954. Pp. xiv, 619.

The jacket heralds this volume as a report of work, designed to test a new theory, performed on animals and man by an interdisciplinary team of psychiatrists, psychologists, physiologists, neurologists, biochemists, and neurosurgeons. "On the basis of . . . experiments with animals, study of schizophrenic patients was begun. In [the patients] there was an apparent physiological abnormality indicated by unusual intensity and duration (spiking) of the brain waves recorded from a key region in the facilitatory circuit . . . mild electrical stimulation [not to be confused with shock therapy] of the facilitatory circuit [was applied] in conscious patients. . . . Some, though not all, of the patients who had had severe forms of schizophrenia for many years responded to such stimulation with marked and continued improvement."

In brief, the stated theoretical framework within which this study was performed considers the schizophrenic process "a basic deficiency on the hedonic level of integration" which is "manifested by defective self-awareness, faulty integration of emotional reactions, and a type of intellectual impairment." In the first chapter, the author compares the schizophrenic state with that seen in "deep reverie or sleep" and suggests the hypothesis that "there must be some inhibition of cortical activity" in the "decompensated" schizophrenic. On the basis of previously reported electrophysiological experiments involving stimulation of various neuronal structures with recording of the enhancement or depression of movements, he suggests that "facilitatory" and "inhibitory" circuits exist in the central nervous system. Specifically it is claimed that "Stimuli to the prefrontal and temporal cortex (those parts which, when ablated in the human, produced lessened affect from painful memories) fire the septal region," and that "Stimuli to the septal region activate the cortex, especially the frontal and temporal lobes." Thus a facilitatory circuit involving the septal region (defined in gross anatomical terms) is postulated. The ensuing animal studies and the stimulations and destructions of the septal region in man grow from this argument.

Whereas the chapter on theoretical concepts is voluminous, the chapters devoted to data are superficial. The first chapter, describing metabolic changes following brain lesions in animals, is one of the best documented; yet even this chapter suffers from inadequate description of the metabolic tests used, and from poor organization of the results which makes difficult an analysis of the data. In the next chapter, which deals with behavioral changes following cerebral surgery, the conditions under which the results were obtained are not described, controls are not mentioned, and there is no evidence that behavioral tests were administered. The chapters on the electrophysiological studies, although in some cases adequate, are not as complete as they would be in a journal devoted to the field (a point borne out by comparing the presentations of the same data where they have been published in both channels). In many instances, whole chapters are devoted to reports of sporadic stimulation of one or two animals and only rarely is the locus of the stimulating electrode represented by anatomical verification. Another criticism of these chapters is that relevant data available in the literature are not systematically reviewed. There are, in fact, some rather surprising omissions: *e.g.* no mention is made of the 'spontaneous' electrical 'spike' activity in basal forebrain structures of presumably 'non-schizophrenic' monkeys, a finding which casts doubt on the premises upon which the entire program is based.

Regarding the studies in man, little more need be said from the standpoint of the scientist than to remark on the absence of any sort of control procedure. From the standpoint of a surgeon, however, it seems incredible that any surgical effort which has not progressed beyond a 25% morbidity and 10% mortality from infection alone should be taken from the laboratory and applied to patients. Have we completely forgotten the first principle of surgery: *Prime non nocere?*

Are these surgical excursions warranted by the therapeutic results? It is clear that even by the psychiatrist's standard of 'marked improvement' (which includes patients classified more conservatively by the psychologist as either slightly improved or deceased because of brain abscess), only a 25% unequivocal therapeutic result is obtained. In the absence of a control group, this value may be compared with an overall improvement rate in schizophrenia which stands around 33%. The therapist is unimpressed and agrees with one of the discussants that perhaps the surgery (with all its problems) is far ahead of the therapy.

As a student, the reviewer is in sympathy with the aims of the book—aims which are unassailable. Nor are the above comments to be taken to indicate that the reviewer can tell anything regarding the possible fruition of the program undertaken. That, of course, is just the point: the book does not help in one way or another. Even the theoretical discussion lacks clarity, which is due for the most part to the naïve dualism which pervades the discussion, and to the lack of operational analysis of the concepts brought forward. Regarding the wonders of a "multidisciplinary approach" this publication reconfirms the reviewer's opinion that only when the disciplines come together within one cranial cavity can such an approach prove fruitful in this stage of scientific endeavor. This volume attests to the fact that psychiatrists do not become scientists by physical contact with scientists, that electrophysiologists make no greater contribution to an understanding of schizophrenia because they eat lunch with psychologists. Instead there is a tendency for each group to show less responsibility for statements made in the realm of the other's discipline, since expert opinion is so easily available.

The intentions were good; the result is not. As a mimeographed progress report with limited distribution, these chapters would have been of use to those actively concerned with the problems treated. Communication would have been enhanced, corrections of communicants' preconceptions would have been offered, and research efforts might have been increased and improved. As a published book, suggesting that a new avenue of approach to schizophrenia has been opened, these same chapters must be described as an atrocity. Much as the Commonwealth Fund is to be commended for supporting this endeavor, it is to be censured for making public in this form preliminary data which would, for the most part, be unacceptable in the more established channels of publication. This disservice is greatest to the team which is being supported. Only in attempting to meet accepted scientific (and therapeutic) standards—or to go beyond them—can knowledge be accumulated.

Institute of Living KARL H. PRIBRAM

Introduction to Factor Analysis. By BENJAMIN FRUCHTER. New York, D. Van Nostrand Company, 1954. Pp. xii, 280.

At long last a textbook has been written that can be used successfully in a beginning graduate course in factor analysis. In writing *Introduction to Factor Analysis,* Fruchter has provided enough material in eleven chapters covering slightly more than 200 pages for a one-semester course of three hours credit. Following Thurstone's point of view, the author has achieved a balanced exposition of the field in the development of the general principles of factor analysis, in the explanation of several methods of factor-extraction and axis-rotation, and in the description of carefully chosen experimental studies that illustrate the numerous potential applications of the techniques. The arrangement of the chapters is sound both from a logical and from a pedagogical standpoint. The exposition is particularly well suited to the needs of the beginner who wishes to have a manual at his finger tips. In step-by-step fashion, a given procedure is outlined and usually illustrated with numerical data, and at the end of nearly every chapter, there are illustrative problems and answers.

The first four chapters may be considered as introductory and fundamental to an understanding of the following seven. After describing in Chapter 1 the general nature of factor analysis, introducing a few preliminary concepts, and discussing Spearman's two-factor theory and Holzinger's bi-factor theory, the author devotes his second chapter to a consideration of cluster analysis, which furnishes the reader with a further understanding of the objectives of factor analysis and introduces him to a preliminary step in some of the grouping methods. In Chapter 3, the essential background in mathematics is succinctly and clearly presented in terms of the minimum requirements of matrix-algebra and matrix-geometry. Ample illustrative material is included. In Chapter 4, attention is given to the two basic assumptions concerning (*1*) the representation of an individual's test-score in terms of common factors through use of the specification-equation and (*2*) the reproduction of correlation coefficients from knowledge of common-factor loadings through use of the fundamental equation of factor analysis. Different types of variance are defined, and application of the fundamental factor theorem is illustrated.

In Chapter 5, the diagonal and centroid methods of factoring are explained and

illustrated in detail. In Chapter 6, multiple-group and principal-axis methods of factoring are described, and applications of these methods of analysis to given data are clearly demonstrated.

After explaining the concepts of positive manifold and simple structure, the author proceeds in Chapter 7 to illustrate the technique of rotation of orthogonal axes both graphically and analytically. One of the high spots of the book is the thorough, step-by-step description of orthogonal rotation of the five axes of a centroid matrix consisting of eleven test-variables. Accompanying each of the eleven rotations is an explanation of the reasons underlying the execution of the particular rotation in question. The only major suggestion that the reviewer would make at this point is that a summary list of guiding principles might have been included to which the reader could refer—a list in which cautions as to what not to do might be summarized in addition to positive statements regarding courses of action to be taken.

A minor suggestion regarding the calculation of the transformation martix might also be in order. It would seem advisable to point out that the entries in each of the successive Λ matrices actually constitute cosines of the angles of rotation between corresponding pairs of axes. In a recent seminar on factor analysis, students with a fair amount of mathematical background considered the finding of direction numbers and normalizing the vectors to be a waste of time when actually one had only to measure the angle between two positions of an axis with a protractor and to look up the cosine of the angular separation in a table. It would appear that a justification of the procedure followed, and mention of the fact that it does lead to the realization of the cosines of the angles between corresponding pairs of axes, would further the reader's understanding of the value of the transformation matrix and of the value of the normalizing process.

The single-plane method of rotation and Harris' approach to oblique rotation based on the use of extended vectors are described in Chapter 8. Additional explanation concerning the interpretation and use of the various sets of plots in relation to the single-plane method would have been helpful. In Chapter 9, an interpretation of the nature of the factors of a rotated factor matrix of eleven test-variables is presented. A list of the items that should be reported in a factor-analytical study is given.

Particularly helpful to the beginner in the interpretation of the results of factor analysis is the material in Chapter 10. The findings in a variety of studies from various content-areas of psychology are reported and discussed in considerable detail. Actually this chapter can be assigned to advantage after study of Chapter 4 or Chapter 7.

In Chapter 11, "Some General Considerations," numerous problems are briefly treated. Although Fruchter ably distinguishes between simple axes (normals to hyperplanes) and primary axes, more might have been said concerning the differences between *structure* and *pattern*. In particular, interrelationships could have been pointed out between both perpendicular and parallelogram-type projections of a test-vector upon the two types of axis. Simultaneously, twelve or fifteen matrix equations could have been written to represent the numerous possible interrelationships occurring among various combinations of the following parameters: factor loadings (two types), correlations of tests with factors, correlations between

factors, and intercorrelations of tests. Amplification of the existing geometric figures could have been effected to make such interrelationships apparent at a visual level.

Other topics that are considered in Chapter 11 include a discussion of oblique versus orthogonal solutions; the current status of factor-identification (the number and types of factor that have been established); and conditions that influence the outcomes of a factorial study, such as the mode of selection of a sample, the background, training, experience, and set of the examinees, the type of correlation coëfficient employed, the existence of experimental and linear dependence, the time-limits allowed, and scoring formula used. In addition, some limited consideration is given to the *Q* and *R*, *O* and *P*, and *S* and *T* techniques. The estimation of factor scores and the use of mechanical aids in computation are briefly described. A closing section on "Next Steps in Factor Analysis" points to the encouraging amount of progress that has been made in the development of analystic solutions since the manuscript for the book was started. A useful section on the "Calculation of the Estimated Factor Loadings of a Variable not Included in the Original Analysis" is included in the *Appendix*. A bibliography of material written primarily since 1940 is another helpful feature.

In all, Fruchter's *Introduction to Factor Analysis* is a valuable contribution to the field of psychometrics. As a teaching aid for a beginning course, it is without peer. In a six-week summer course in factor analysis at the University of California the book worked out exceedingly well. As a manual which the research worker with limited mathematical training can use in the diversified content-areas of the social sciences, the book also will serve a valuable purpose. Every student of factor analysis and every research worker who has occasion to employ factor-analytical techniques should have a copy.

University of Southern California WILLIAM B. MICHAEL

The Clinical Interaction. By SEYMOUR B. SARASON. New York, Harper & Bros., 1954. Pp. x, 425.

This book may be evaluated from two different points of view, and one's final decision must be determined very largely by one's decision as to the proper frame of reference. On the one hand, we may accept it as a contribution to the Rorschach literature, in respect to which we agree to take a large number of things for granted (such as the validity of the Rorschach Test as a whole, the diagnostic usefulness of psychotherapy, the truth of certain psychoanalytic doctrines, and so forth). On the other hand, we may ask ourselves questions regarding the contribution of this book to the science of psychology, using the term in a reasonably rigorous sense. The fact that these two universes of discourse exist and require separate consideration is in itself an interesting comment on the state of modern psychology.

Judging the book then in the first place as a contribution to the Rorschach literature, we should say that it is clear, well written, knowledgeable, critical, and altogether superior to the great majority of books in this field. It is written in three parts, but throughout it is based on what the author calls "the clinical interaction." "This book stems from the belief that practically all clinical problems are concerned with data obtained from an interpersonal interaction and that regard-

less of how interactions differ certain variables are always operative and must be taken into account." This means that as far as the Rorschach is concerned, neither the derivation of scores, nor the consideration of the relationships obtaining between scores is enough; the crucial part of the clinician's task must be the ascertainment of just what goes on between himself and *S*, how *S* formulates the task to himself, how this "clinical interaction" determines the results, and so forth.

The first part of the book, dealing with "situational variables," puts this point of view at some length; the second part of the book discusses a whole body of Rorschach literature in a critical manner, emphasizing the factual basis for common clinical judgments; the third part, entitled "Individual Interpretation," presents in some detail six cases which illustrate the points made by the author.

With much of what the author says it would be difficult to disagree. There is stress on experimental verification throughout. "When interpreting the record of an individual it is the obligation of the clinician to avoid using assumptions which, although stated with frequency and authority in the textbooks, have little or no basis in the research literature. The fact that many clinicians are not researchers does not absolve them from critically evaluating the research literature on the instruments they employ in practice. We make these statements because of our impression that many clinicians utilize the Rorschach in a way that suggests an uncritical acceptance of traditional assumptions." Sarason's critical survey of the literature suggests to him a certain minimal number of statements which may usefully and validly be made about the Rorschach. These statements and this discussion certainly are much more down to earth and realistic than the type of material one has been used to in connection with the Rorschach Test.

In spite of these good points, the book appears less valuable than it might otherwise have been because the author fails to take into account sufficiently the second of the two universes of discourse mentioned at the beginning of this review. He seems to be aware of this omission because he points out to "the reader who is not a clinician . . . that the clinician is expected to give some kind of answer to the problems which are presented to him, and he cannot avoid answering the problems because of uncertainties about the assumptions upon which he operates" (p. 295). To what extent this social pressure should lead the clinician to make use of unvalidated tests, inspired guess work, and the like, this is not the place to discuss. From the point of view of the science of psychology, it must be pointed out, however, that a contribution has to satisfy much higher standards before it can be regarded as acceptable. Sarason himself draws attention to the fact that different authors use different instructions and different procedures, different scoring systems, and different principles of interpretation; he reports investigations showing that the personality of *E*, as well as the setting, the social class of *S*, and many other variables, affect the outcome of the test. Nevertheless, he spends a good deal of the time comparing different researches, which differ in all these respects, and it is difficult to see just what the point of such a procedure can be. Surely, before any results can be compared, the effects of variation in procedure and other matters must be known and discounted. This is only one example of a split which seems to go through the book. Criticism of individual experiments and comparison of their results is carried out with insight and in terms which would seem reassuring to the scientifically trained investigator, yet throughout there is a failure to face

the fact that major requirements for the comparison of individual studies, such as those outlined above, are lacking. The critical reader, who follows all the experimental documentation regarding the many uncontrolled factors determining Rorschach responses, may end up by asking whether knowledge regarding this test has really reached a point where anything positive whatsoever can be said regarding it. He may also wonder why this test was chosen in preference to other methods of investigating personality which are less subject to such damaging criticism.

There are a few minor points of criticism which should be mentioned. It is not always clear why the author quotes certain studies rather than others; thus, he relies heavily on Wittenborn's factorial analysis of the Rorschach Test, but gives no reference to the analyses by Sen, Cox, and others, which would superficially seem to be even more relevant. There is some evidence of poor proof-reading. Thus, the reviewer found his name (incorrectly spelled) in the author index; looking up the page given there, he found no mention of his work until two pages later, where his name occurred, again incorrectly spelled. A rapid survey disclosed that all references to the bibliography were misplaced by two pages, an error which might be very annoying to the unwary reader.

Maudsley Hospital H. J. EYSENCK

Sociological Studies in Scale Analysis. By MATILDA WHITE RILEY, JOHN W. RILEY, JR., and JACKSON TOBY. New Brunswick, Rutgers University Press, 1954. Pp. xii, 433.

The purpose of this book is admirably described in the preface. "This book is written for researchers using qualitative data both in the behavioral sciences and in such applied fields as market research and public opinion polling. It shows empirical examples of scales, and sets down in detail with actual worksheets the operations involved. It examines into several of the theoretical implications and assumptions of social measurement and sets forth specifically some operating procedures. Primarily, however, it is written for sociologists, in that it suggests how such tools as scaling may be used for theoretical as well as substantive clarification." The reader is warned that "While every effort has been made to keep the exposition clear, this book is not to be *read* in the usual sense. There can be no ready simplification of the essentially difficult and complex concepts and procedures involved. It is a book to be worked with." The authors are undoubtedly right in saying that this book should not be read in the usual sense. It is doubtful to the reviewer if it should really be called a book at all in the ordinary sense of the term. It is a collection of articles, notes, and theoretical discussions, many mere reprints of published papers thrown together in a rather haphazard fashion and without any very definite plan. Thus, for instance, after 272 pages of frequently very complex argument and sophisticated statistical discussion dealing with the application of Guttman's scales to sociometric problems, the reader is suddenly treated to a very simplified, told-to-the-children kind of introduction to the Guttman scaling techniques. There seems to be no reason for this odd placement, except that this introduction happened to occur in a published paper which, for other reasons, was thought to be more relevant to a later part of the book than to an earlier one.

Unless the reader, therefore, is relatively knowledgeable and sophisticated to begin with, he will almost certainly be completely puzzled by the contents of this volume in their present arrangement. The expert will undoubtedly find a number

of interesting leads, methods of analysis, and new areas of application; he may also be grateful to have reprinted in easily accessible form a number of important papers which might otherwise not be readily obtainable. Even so, he might have preferred a book of the more usual kind, *i.e.* one written with a clear plan and purpose, with a beginning, a middle, and an end, and with an arrangement both logical and sensible.

Another difficulty which the psychological reader may find with this book is the terrible prolixity which seems to afflict the writers, particularly in connection with theoretical subjects. An attempt seems to be made to relate sociological theory to scaling, but this attempt is very largely on a semantic level, which is perhaps not surprising since sociological theories, particularly those connected with Parsons and Shils, are not usually of a kind which lend themselves to verification or disproof. An illustration of the perfunctory way in which the task is performed may be found in a footnote on page 39. The authors are talking about an analysis of the rôles of individual players and their structure, or interrelationship. They go on to refer to "the distinction between *action* patterns and *plurality* patterns in Leopold von Wiese and Howard Becker, *Systematic Sociology,* John Wiley and Sons, 1932. Although the present empirical effort, by contrast, focuses only upon limited aspects of patterned acts, on the one hand, and structured rôles, on the other, there may well be an element of theoretical continuity in the two frames of reference. Nor is this proposal unrelated to such important distinctions as are made in the theoretical works of Pitirim Sorokin and Florian Znaniecki." Such phrases as "there may well be" and "nor is this proposal unrelated" indicate that the integration between theory and what is done in this book is of a very loose and woolly kind. The example quoted illustrates a tendency which the reader will find throughout the whole book.

A further point which will occur to the critical reader is this: Technical procedures in scale analysis are discussed in great detail, and working methods set out sufficiently well to enable anyone to follow them easily. What seems to be lacking, however, is the much more important discussion of fundamental points of criticism. The reader of this volume will in vain seek enlightenment as to the grounds on which these scaling techniques have been attacked, and equally vainly will he seek for a refutation of these criticisms.

Altogether then, this appears to be a hastily put together compendium dealing with a large number of loosely related points in the general field of scale analysis, touching in a very non-rigorous fashion on a great variety of sociological theories of the more speculative and semantic type. It has much of interest and importance to say to the expert, but would prove a very purgatory for the beginner trying to find out about scale analysis. The book is enlivened by a number of empirical studies which appear to be essentially of a trivial nature, and are not designed to answer any very obvious psychological questions.

Maudsley Hospital H. J. EYSENCK

Child Psychology. By ARTHUR T. JERSILD. Fourth edition. New York, Prentice Hall, 1954. Pp. xii, 676.

It is a pleasure to review this completely overhauled fourth edition of Jersild, which combines a very conversational, almost sprightly, style with an affective understanding of children and a thorough command of the empirical literature.

It is up to date. It is integrated. It has warmth and feeling without sentimentality. One's impression is that Jersild knows children as well as about children. His delight in them and sympathy for them will charm even the most crusty Watsonian reflex-counter.

Child Psychology will be a tremendous success with undergraduates, parents, and all others who want to learn about *children*. Hard-hearted academicians, those of us who view children as just so many *S*s with somewhat different characteristics than rats and sophomores, will find it *too* child-centered. Some traditional questions are omitted entirely (*e.g.* form-color dominance, 'relational' versus 'absolute' learning). Others are approached primarily in terms of their common-sense, everyday, practical implications (*e.g.* orality, affection, nature-nurture).

The differences between this edition and the third are enormous. Most impressive to the reviewer are these: a vastly enlivened style; a point of view emphasizing the 'self' and its hygiene without any recondite, esoteric lingo; friendly references to such previously black-balled thinkers as Freud, Horney, Sullivan, Kierkegaard, and Moreno; a drastic reduction in the number of tables and charts that gave much-too-much detail about many-too-many investigations; and Jersild's willingness to venture an opinion rather than hide behind the data of half-a-dozen partly contradictory investigations. No longer do we have a vast string of studies listed one after the other with the *N*s, means, sigmas, and the grade-placements of the *S*s. Instead, we have the results of one study related to the results of others, and even some tentative conclusions.

Another advantage of this edition is that, while up to date, it is not positively encyclopedic. Jersild has selected his material, which is as it should be; he leaves the instructor something to do. Some will add comments based on Mowrer and on Miller to the discussions of fear and anxiety. Many will deal with the development of intelligence at much greater length. Others will refer in detail to the Wickens' paper on neonatal conditioning, which is only cited in a footnote, and to Spelt's paper on prenatal conditioning which is not referred to at all.

In short, Jersild emphasizes children, their problems and their happiness. Yet he manages to give a high proportion of the developmental data as well. A typical illustration of this difficult achievement is his section on competition. It is recommended as a representative sample for instructors who are considering a change of text.

University of California JOHN P. MCKEE

Time Distortion in Hypnosis: An Experimental and Clinical Investigation. By LINN F. COOPER and MILTON H. ERICKSON. Baltimore, Williams & Wilkins Company, 1954. Pp. ix, 191.

This book, produced by two physicians, contains an experimental section written by Cooper and a clinical section written by Erickson. The book is both provocative and bizarre—provocative in that the problem is one that deserves more thorough investigation, and bizarre in that the experimental design is haphazard. The authors are concerned with distortion (lengthening) in apparent time which depends on, or is said to depend on, mood, drugs, crises, dream-states, and so forth. The book is concerned with a greater degree of distortion than is usually encountered, although the actual extent is never specified. The procedure employed has been described as bizarre for a number of reasons: many experiments are performed but no effort

is made to validate any one conclusion, controls which are not controls are utilized, practice-effect is frequently ignored, treatment of the data is oversimplified, and so on.

When the reality of time-distortion is evaluated by the use of mathematical problems, or by *S*'s hearing (in his hallucinated world) a lengthy discussion of a disease (that he is not conversant with in his waking state), there is failure to confirm the theory. In the face of this failure, the authors fall back on the reported "real nature of the experience" and attempt to explain it (away?) by stating that the answer must lie in the person's past experience or in "new combinations thereof."

The psychologist is wary of the testimony of men of "unimpeachable integrity," not because he believes that deliberate falsification is involved, but because other and generally simpler explanations of a valid nature are possible. In this book it seems clear that the investigators conveyed to their *Ss* (in both the training and experimental sessions) their own beliefs concerning the reality of time-distortion. Inasmuch as there was, and could be, no concurrent reporting by the *Ss* in their time-distorted state, but only elaboration *afterwards,* it is very possible that the *Ss* were reacting to an implied posthypnotic suggestion and, for this reason, gave the desired results. Posthypnotic suggestion rather than time-distortion can explain most, if not all, of the data in the experimental section.

Erickson's cases, while brief, are interesting and suggestive but do not confirm, directly or indirectly, the reality of time-distortion. With regard to applicability, a question arises. Probably less than half of the general population enter into a moderately deep hypnotic state within a feasible period of time, and of this group only a certain proportion (not specified) can be trained to show time-distortion.

This reviewer believes that Erickson and Cooper have described nothing more than a modification of a posthypnnotic technique. It may possibly be said that if hypnoanalysis is a shortcut for conventional analysis, then so-called time-distortion elicited in hypnosis is a shortcut in hypnoanalysis.

F. L. MARCUSE
The State College of Washington

Methods of Research: Educational, Psychological, Sociological. By CARTER V. GOOD and DOUGLAS E. SCATES. New York, Appleton-Century-Crofts, 1954. Pp. xx, 920.

According to the authors of this large work on research methods, it is intended for field workers, graduate students, and advanced undergraduates in education, psychology, and sociology, "who would evaluate the quality of conclusions either as producers or consumers of research." The apparent purpose of the authors is to give students an awareness of the characteristics and uses of the various experimental approaches to problems in these fields. The attempt to fulfill this purpose is made in a series of chapters which deal with such topics as the justification for research, location of research problems, library techniques, the historical method, descriptive methods, the experimental method, case study methods, developmental approaches, and reporting of research.

The chapters vary in difficulty from some which are quite elementary how-to-do-it chapters, such as the one on location of research problems, (chapter II), to some which are quite advanced, such as a sub-chapter called "Descriptive Method:

Analysis" (chapter V), which contains a detailed discussion of the assumptions and purposes of descriptive analysis. The former seems directed to persons who have been told to find a research problem and who do not know where to begin; the latter to persons concerned with the basic problems of a particular approach. There is also an unevenness in the authority which the writers seem to bring to the various topics, ranging from the full and probably valuable treatment of description and descriptive techniques in chapters V and VI to the very inadequate treatment of experimental method (chapter VII). Perhaps the authors' expertness and interest in these topics might be indicated by a comparison of the number of pages in chapters V and VI with the number in chapter VII: chapter V (in three parts) approximately 290 pp.; chapter VI (in three parts) 140 pp.; chapter VII, 31 pp.

Since for many psychologists, the most important chapter in a book such as this is the one on experimental methods, a couple of examples of what appear to be serious shortcomings in that chapter will be mentioned. First, a frequent procedure is to list a series of terms with no indication of their meaning and sometimes in connection with misleading statements. For example, in a section on illustrative classifications of experimental designs, there is a one-sentence discussion of logical classifications in which Mill's canons are listed but not discussed because, it is held, they are not serviceable in modern experimental design. Second, there is no examination of the relationship between hypothesis and experiment. Certainly anyone hoping to evaluate the quality of conclusions in contemporary psychology would require much instruction on this problem. It is doubtful that experimentally oriented readers will find this book of value. Those interested in survey and description may well find some stimulating sections.

University of California

DONALD A. RILEY

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXVIII JUNE, 1955 No. 2

THE EFFECT OF ATTITUDE UPON THE PERCEPTION OF SIZE

By ALBERTA S. GILINSKY, Columbia University

In the absence of other indicants to distance, the size of the retinal image necessarily determines our perceptions of two dimensions, width and height.[1] Holway and Boring have shown that as increasing sensory data are made available, perceived size comes to depend less on retinal size and more on object-size as distance is varied.[2] For free binocular regard, functions relating the adjusted size of a comparison-object to the distance of a standard object were close to or exceeded the function for size-constancy, at least up to 120 ft., the maximal distance used in their experiment. What happens at greater distances? Two experiments have been reported which match a relatively near-by object to a distant object: one, a stake, half-a-mile away; the other, the moon, 239,000 miles away.

Under open-air conditions favorable for viewing distance, Gibson required his *Ss* to estimate the real, correct height of wooden stakes, from 14 to 784 yd. (nearly half-a-mile), in terms of a numbered series of comparison stakes, 14 yd. away.[3] Five practice trials were given with the test-objects at the same distance as the comparison objects, 14 yd., followed by 25 trials at each of 6 distances, beginning with

* Accepted for publication April 9, 1954. This research, conducted under terms of USAF Contract No. AF 18(600)-196, was administered by the Skill Components Research Laboratory, Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas. The basic problem was formulated by Professor Edwin G. Boring to whom the writer is indebted for many valuable suggestions.

[1] William Lichten and Susan Lurie, A new technique for the study of perceived size, this JOURNAL, 63, 1950, 280-282; A. H. Hastorf and K. S. Way, Apparent size with and without distance cues, *J. Gen. Psychol.*, 47, 1952, 181-188.

[2] A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37.

[3] J. J. Gibson, Motion picture testing and research report No. 7, *Army Air Force Aviation and Research Reports*, 1947, 200-211.

the farthest distance and progressing toward the nearest. The results for the 71-in. stake were typical. The mean estimate at 14 yd. was 71.9 in. ($SD = 1.8$). At 784 yd., with the intervening ground continuously visible, the 71-in. stake gave a mean estimate of 74.9 in. ($SD = 9.8$). Although variability increased with distance, Gibson concluded that "an object can apparently be seen with approximately its true size *as long as it can be seen at all*."[4]

Boring, on the contrary, found that the apparent size of the full moon on the horizon was equalled by an 8.5-in. disk 12 ft. away—that is, the moon on the horizon was matched by a disk subtending an angle of 3° although the moon itself subtended an angle of only 0.5°.[5] Boring summarizes his observations as follows: "It is impossible to perceive the moon as big as it really is (2160 miles across) or as small as its retinal image is (0.5° across). You see something in between, nearer retinal size than object size."[6]

These experiments leave us an unexplored range (between a half-a-mile and astronomical distances) and some very perplexing problems. Their results are contradictory: one indicates that size-constancy fails at great distances; the other that it does not. Boring accepts both conclusions and compares the paradox to the dilemma of the railroad tracks, now seen to converge, now seen not to converge.[7] He suggests that we may be dealing with two systems of perception, two observational attitudes corresponding perhaps to Gibson's distinction between the visual field and the visual world.[8] The problem, then, is to specify these two types of perception, this field and this world, in operational terms. How can this difference in observation be induced? If it depends upon different conditions of the organism and not upon different conditions of stimulation, then what are the alternative modes of response to the same stimulation, and how are they related to distance?

The present study attempts to answer these questions by investigating the perception of size of objects under two contrasting observational sets: one for matching 'objective' size; and the other for matching 'retinal' or 'projected' size.

Apparatus and Procedure

To provide a sufficient range and maximal opportunity for the object-directed set to exert an influence, the experiment was conducted out of doors in daylight with all the usual cues of distance available. The experimental site consisted of a fairly level stretch of grassy terrain parallel to an inactive airport runway, 5,000 ft. long.

The general plan of the experiment required that *S* match the size of a standard

---

[4] Gibson, *The Perception of the Visual World*, 1950, 186.
[5] Boring, The moon illusion, *Amer. J. Physics*, 11, 1943, 55-60.
[6] Boring, Visual perception as invariance, *Psychol. Rev.*, 59, 1952, 146.
[7] *Ibid.*, 142 ff.
[8] Gibson, *The Perception of the Visual World*, 26-43.

stimulus-object placed at various distances directly ahead of him, by altering the size of a variable stimulus-object, 100 ft. away and 36° 26′ to the right of the direct line of regard.[9]

Stimulus-objects. The stimulus-objects were plane white isosceles triangles, constructed of sheet aluminum and placed perpendicularly. They were seen against a background of grassy terrain and remote trees and buildings at the far end of the

FIG. 1. SCENE AS VIEWED BY *S*
The 66-in. standard triangle at 200 ft.

field. Fig. 1, a photograph taken from *S*'s station, shows one of the standard triangles at 200 ft. The surface of the ground, as may be noted, is clearly visible and provides cues of perspective and texture. Other monocular cues were provided by aerial perspective, light and shade, and head-movement parallax. Binocular vision was employed throughout the study, thus stereoscopic vision may have supplied cues at the shorter if not at larger distances.

(*a*) *Standard.* Four standard triangles were used, which ranged in 12-in. steps

[9] R. B. Joynson (The problem of size and distance, *Quart. J. Exper. Psychol.*, 1, 1949, 119-135) found that the effect of size-constancy was maximal at approximately this degree of separation.

in base and altitude (which were equal) from 42 to 78 in., and in area from 882 to 3042 sq. in. In these triangles, a leg is 1.12 times the altitude or the base. The standard triangles were placed, one at a time, at 6 distances: 100, 200, 400, 800, 1600, and 4000 ft. from *S*.

(*b*) *Variable.* The variable triangle was identical in shape and color with the standard triangles. It could be varied in size by raising or lowering it into a pit in the ground. An attempt was made to provide a range in size-variation which exceeded the range in altitude of the standard objects at both ends. When elevated to maximal height, its altitude measured from ground level was 7 ft. 2 in. (see Fig. 2). While this upper limit exceeded by 8 in. the altitude of the largest test-object, it proved to be insufficient, as will appear later. The variable triangle could be lowered

FIG. 2. VARIABLE TRIANGLE AT MAXIMAL SIZE
The man beside the triangle gives a cue, unavailable to *S*, of relative size.

below ground level to provide a zero value of the size-range. Since the nature of the mechanism was such that settings of the triangle at an altitude of only a few inches could not be accomplished as precisely as settings at greater altitudes, the variability of the settings at the lower end of the size-range was correspondingly greater.

The adjustment of the variable triangle was under the remote control of *S*. Compressed air was used to operate the lifting apparatus, as shown schematically

in Fig. 3.[10] The apparatus was contained by a T-shaped pit, 8 ft. 5 in. deep, lined with cinder blocks. A side-view is shown at A. The triangle was supported by a double steel beam, 16 ft. in length, with a hinged coupling on the back of the triangle approximately 4 ft. from its top. The opposite end of the beam was connected to the end of the piston shaft of the air cylinder C by a hinged coupling at E. The beam supporting the triangle was itself supported by a second beam 8 ft. in length with a hinged coupling at B, exactly in the center of the longer beam. The 8-ft. beam was hinged to the base of the apparatus directly beneath the triangle.

This mechanical linkage resulted in the translation of horizontal motion of the shaft of the air cylinder into vertical motion of the triangle in a frontal plane.

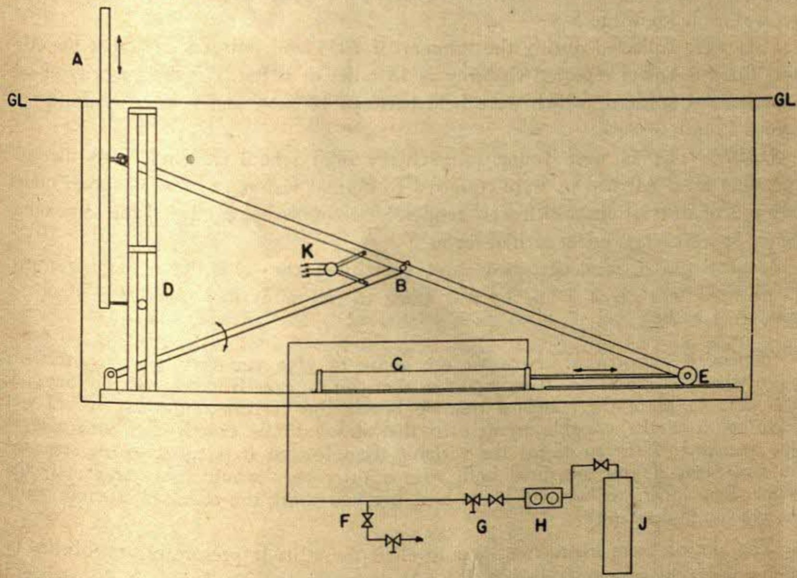


FIG. 3. DIAGRAM OF APPARATUS FOR VARIABLE TRIANGLE

By appropriate adjustment of the input valve G in the line from the high pressure storage tank J and the exhaust valve F, the triangle could be raised or lowered to any desired position and maintained at that position. Two small wheels supported by short metal extensions from the base of the triangle traveled in vertical guide rails D located directly behind the triangle, and eliminated wobbling.

The altitude of the variable triangle was recorded from a remote electrical indicator at *E*'s station.[11] A rotary potentiometer was attached at K to the supporting structure of the triangle. Any change in the angle of intersection of these two members resulted in a change in the position of the potentiometer shaft and a consequent variation in resistance between the variable terminal of the potentiometer and either end terminal. The altitude of the triangle was a sine function of half the voltage drop in the circuit, a drop that was indicated by a milliameter in series with a 1000-ohm resistor. Meter readings were converted to triangle alti-

[10] The apparatus was designed and built by Mr. Edward Palasthy, Department of Mechanical Engineering, Columbia University.

[11] The method was devised by Dr. John Lott Brown, Department of Psychology, Columbia University.

tude from a calibration curve that was determined empirically. The range of scale-readings representing the change from minimal to maximal altitude was adjusted by means of a variable resistor in series with the potentiometer. The method was rapid and convenient and it avoided cues which might have arisen from *E*'s approach and measurement of the variable stimulus-object.

Throughout the study an attempt was made to avoid giving *S* any adventitious cues of distance. Comparison of the stimulus-objects with familiar objects was prevented as far as possible. The standard triangles, for example, were not seen by *S* as they were moved to the different stimulus-points and neither their size nor number was made known to *S*.

Data were collected during the summer of 1953 and only on days that the Airport Control Tower reported visibility as 13 miles or better. For the majority of the experimental sessions, which were held between 10 A.M. and 5 P.M., visibility exceeded 15 miles.

Subjects. The *S*s were young men, chiefly high school students, from the surrounding area. All had or were corrected to normal vision. All the *S*s served under two sets of instructions which were randomly so counterbalanced that the same number of *S*s served first under each of them.

Instructions. A mimeographed sheet containing the 'objective' or the 'retinal' instructions was given *S* and he was asked to follow as they were read aloud by *E*.[12]

'Objective' instructions. Now we are going to give you very specific directions as to what you are to do. It is important that you do exactly as we tell you. Suppose we were to place the standard triangle beside the variable; how big would you have to make the variable triangle so that it would be exactly the same size as the standard? Now so adjust the variable triangle that it is equal to the standard in size—that if you measured both with a ruler they would measure exactly the same. Remember, we wish to know how big you think the standard triangle really is. Do you understand?

The aim of these instructions was to elicit the attitude presumably established in Gibson's *S*s by emphasizing the real, physical, tape-measured size of the standard and variable objects.[13] The task required *S* to look back and forth between the standard and variable stimulus-objects and so to adjust the variable that the two were judged objectively equal.

'Retinal' instructions. Now we are going to give you very specific directions as to what you are to do. It is important that you do exactly as we tell you. As you know, the further away an object is from you the smaller it appears. The moon and the stars, thousands of miles away, look very tiny but we know that they are actually very large. Now, if you were to see a triangle very far away, it would also look pretty small. The question is, how small does it look when it is far away out there in the field? Imagine that the field of view is a scene in a picture or photograph. Every image in the picture is fixed in size. If you were to cut out

[12] The influence of instructions is well attested by the following studies: B. E. Holaday, Die Grőssenkonstanz der Sehdinge bei Variation der inneren und äusseren Wahrnehmungsbedingungen, *Arch. f. d. ges. Psychol.*, 88, 1933, 419-486; T. M. Martin and R. W. Pickford, The effect of veiling glare on apparent size relations, *Brit. J. Psychol.*, 29, 1938, 92-103; M. R. Sheehan, A study of individual consistency in phenomenal constancy, *Arch. Psychol.*, 31, 1938 (No. 222), 1-95.

[13] Gibson, Motion picture testing, *op. cit.*, 203.

the fixed image of the standard triangle and paste it on the image of the variable triangle, would the two images be just the same size? Now, so set the variable triangle that the cut-out image of the standard triangle would be exactly equal to it in size—that the two images would actually coincide.

These instructions were directed to generating the attitude presumably taken by Boring's *Ss* in his experiments on the moon,[14] and by Boring himself in the Holway-Boring study.[15] Here the instructions demanded that *S* look at the standard triangle, fix its image in memory, look at the variable triangle, lay the memory trace of the standard over it, judge too large or too small and adjust. He could look back and forth and check until he could no longer notice a difference between them.

Procedure. After reading the instructions applicable to the day's task, *S* sat in a chair facing the standard triangle—previously placed at one of the six stimulus-distances—and was shown how to operate the apparatus which adjusted the size of the variable triangle. Several practice trials were given to familiarize him with the method.

The standards at different distances were presented in random order, except the nearest distance (100 ft.) was never used first. At every stimulus-distance, the standards to be used during that session were presented, one by one, in haphazard order.

S's adjustments of the variable triangle alternated between 'larger to equality' and 'smaller to equality.' In half of the trials he began with 'larger' and in half with 'smaller.' In both cases he was permitted to make fine adjustments up and down until he was satisfied with the match.

For every standard at every stimulus-distance, from 32 to 36 *Ss* served. The results of our *Ss* were discarded because these *Ss* were unable, under 'objective' instructions, to make the variable triangle large enough, due to the limitations of the apparatus, to bring it to the judged size of the larger standards.

RESULTS

Since an analysis of the data of the individual *Ss* revealed that they gave practically identical results under like conditions (instructions, object-size, and object-distance), the data for the *Ss* were averaged. The means and *SDs* of their settings are shown in Table I. The upper half of the table contains the results under the 'objective' instruction; the lower half, the results under the 'retinal' instruction. The mean size-matches as a function of distance are plotted for the two instructions in Figs. 4 A-D. Fig. 4 A shows the functional relation when the 42-in. standard triangle was used; Fig. 4 B, when the 54-in. standard was used; Fig. 4 C, the 66-in. standard; and Fig. 4 D, the 78-in. standard.

Consider first the curve which represents 'objective' instructions. The variable shows a tendency to *increase* with distance with this observational attitude. For three of the four standards the data for this condition

---

[14] Boring, The moon illusion, *op. cit.;* personal correspondence.
[15] Holway and Boring, *op. cit.*, 25-26.

exceed the function for size-constancy. In the fourth case, that of the largest (78-in.) standard, the results are erroneously small because limited by the maximal size (86 in.) of the variable object. It will be recalled that data obtained from six *Ss* were discarded because these

TABLE I

MEANS AND STANDARD DEVIATIONS (IN INCHES) OF MATCHES OF THE VARIABLE TO EACH STANDARD AT VARIOUS DISTANCES AND DIFFERENT INSTRUCTIONS

Instructions	Standard (in.)	No. of Ss	Distance of standard (ft.)					
			100	200	400	800	1600	4000
'Objective'	42	32	44.7	48.4	56.6	54.6	58.1	56.3
			3.4	6.9	13.1	12.8	11.8	14.3
	54	36	55.3	59.0	62.7	65.2	65.4	69.5
			3.8	8.3	9.8	12.4	12.0	11.6
	66	35	67.8	70.6	71.6	72.1	74.0	75.6
			4.8	8.4	9.5	10.8	9.9	7.3
	78	32	77.6	77.2	76.8	68.7	75.4	78.4
			4.3	6.2	8.2	13.8	9.8	7.7
'Retinal'	42	32	41.0	32.2	23.5	14.9	8.6	4.4
			4.1	4.4	5.6	6.2	5.5	2.8
	54	36	52.6	40.8	30.2	16.7	9.3	4.9
			3.3	6.3	8.1	6.8	5.5	3.1
	66	35	63.8	52.3	38.1	23.5	14.1	6.2
			4.3	7.2	10.5	9.2	7.7	4.3
	78	32	76.4	57.9	44.6	26.7	18.0	8.9
			5.1	7.3	12.7	10.3	9.7	6.0

Ss reported that the variable triangle could not be made large enough to match the size of the 78-in. standard. Presumably, some of the remaining *Ss*, while apparently satisfied with comparison settings at or near the upper limit of the variable triangle, might have produced larger size matches if these had been available to them.

The data for the 'retinal' instructions yield a function for perceived size *decreasing* with increasing distance. The curves drawn through the plotted points for this condition lie within the functions for size constancy and the visual angle, and are fitted by a theoretial equation for perceived size derived by the writer in a previous paper.[16]

[16] A. S. Gilinsky, Perceived size and distance in visual space, *Psychol. Rev.*, 58, 1951, 460-482.

The four curves for 'retinal' matching of Fig. 4 have been replotted in a single figure to facilitate their comparison. Fig. 5 exhibits these functions as a family of curves, with size of the standard as the parameter. The lines drawn through the four sets of data are given by the following single equation:

$$St_c/St = (A + \delta)/(A + D) = (300 + 90)/(300 + \mathrm{D}) \quad \ldots [1]$$

where St_c is the adjusted size of the variable triangle, St is the size of the standard, and D is the distance of the standard from S. The meanings

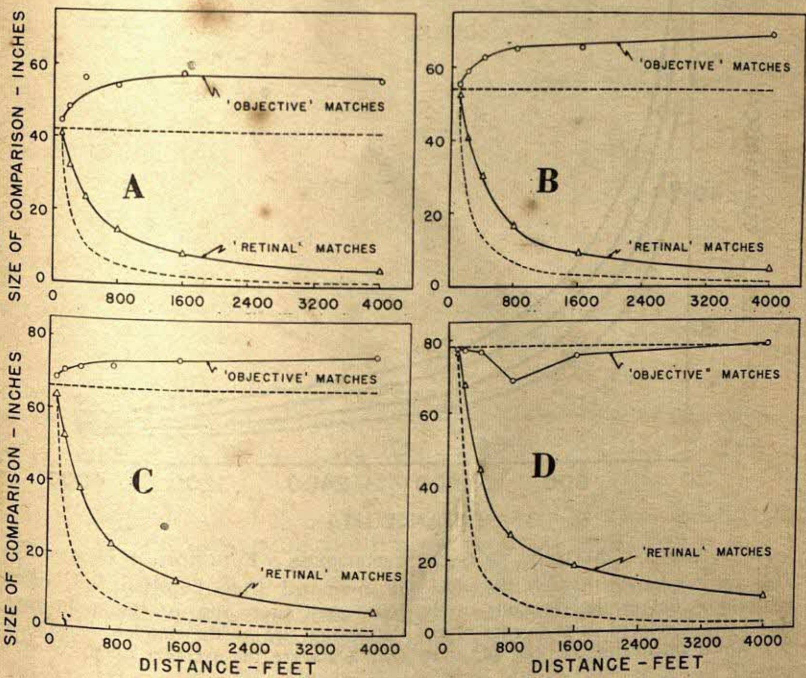


FIG. 4. SIZE OF VARIABLE EQUATED TO EACH OF THE FOUR STANDARDS AS A FUNCTION OF DISTANCE AND INSTRUCTION

A is the 42-in. standard; B, the 54-in.; C, the 66-in.; and D, the 78-in. standard. The horizontal broken line represents the object-size as constant and the bottom broken curve, the actual size of the retinal image or visual angle.

of A and δ are discussed at length in the reference cited. It is apparent that the parameter of size does not affect the form of the function since a single equation provides an adequate fit to all four of the curves.

It may be shown that the ratio of the adjusted size of the variable

triangle to the size of the standard (St_c/St) is independent of the size of the standard. Fig. 6 shows the ratios obtained by dividing the average size of the variable triangle by the size of the standard to which the comparison was equated.

The lower solid curve is the curve of Equation [1] drawn through the data for all standards and all Ss combined for the 'retinal' matching. Each

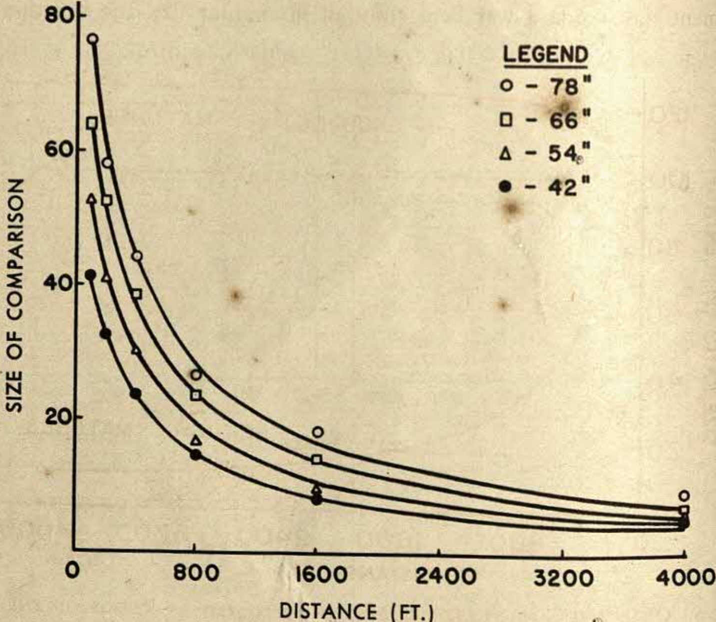


FIG. 5. 'RETINAL' MATCHES TO EACH STANDARD AS A FUNCTION OF DISTANCE
The curves drawn through the data are computed from Equation [1], using identical values of the constants in the four cases. From top to bottom the four curves are displaced in the order of the decreasing size of the standard objects to which they refer, 78, 66, 54, and 42 in. Data are the averages of all Ss.

plotted point (triangle) thus represents the mean of 135 observations. The upper solid curve is drawn through the data for 'objective' matching for all except the largest standard and thus each point (circle) on this function is based on 103 determinations. The results for the largest (78-in.) standard were omitted from this composite because they deviate from the rest of the results obtained under the 'objective' instructions, and are believed to reflect an artificial 'cut-tail' error imposed by the upper limit of size of the variable triangle. To have included them in determining the average 'objective' matches would have led to a false

general impression of greater objective accuracy than was actually achieved. The data summarized in Fig. 6 represent the typical findings of the present experiment.

While this is a simple arrangement of the data, the functions do not appear to resemble those obtained by other investigators. Fig. 7 is a different plot of the data in the form made familiar by Holway and Boring and facilitates comparison with their results.[17] Since in their experiment the standard was kept constant in angular size, the function for

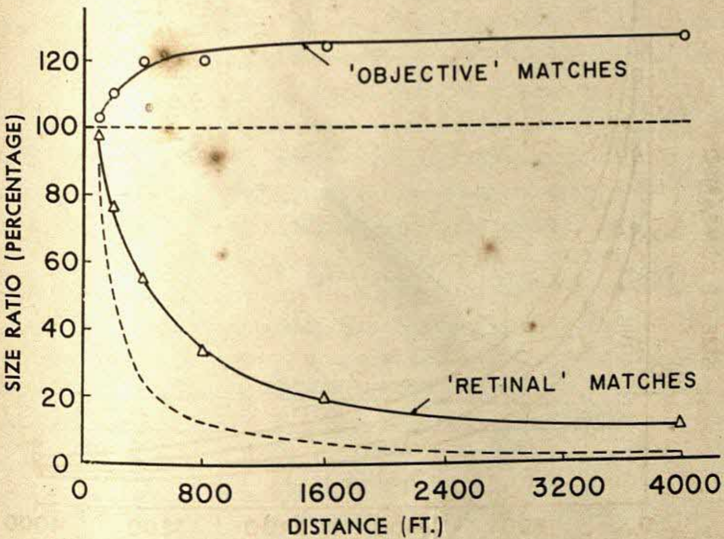


FIG. 6. 'OBJECTIVE' AND 'RETINAL' MATCHES PLOTTED AS RATIOS OF THE SIZE OF THE VARIABLE TO THE SIZE OF THE STANDARD TO WHICH COMPARISON WAS MADE

The points plotted are for all *Ss* and all standards except the largest. Constant object-size and the size of the retinal image (visual angle) are shown by the upper and lower broken lines respectively.

constant object-size is a straight line which rises in proportion to the distance and the function for constant retinal size is a horizontal straight line.

Fig. 7 uses the data of Fig. 6 to show how the size of a triangle would appear to change if its angular size and not its linear size were kept constant as it receded 4000 ft. from *S*. Taking the value for 100 ft. as the base, the means of the size-matches were multiplied by the corresponding distances, by 2 for 200 ft., by 4 for 400 ft., . . . by 40 for 4000 ft. The

[17] Holway and Boring, *op. cit.*, 23 ff.

mean standard size, a theoretical test-object of 60 in. at 100 ft., was also multiplied by distance to illustrate the ideal function for a variable object increasing with the standard object and remaining equal to it. Such an object would subtend a constant visual angle of 2° 52′. Fig. 7 makes the assumption that the relations of Table I would hold for very much larger test-objects, that it is the ratio of sizes being compared and

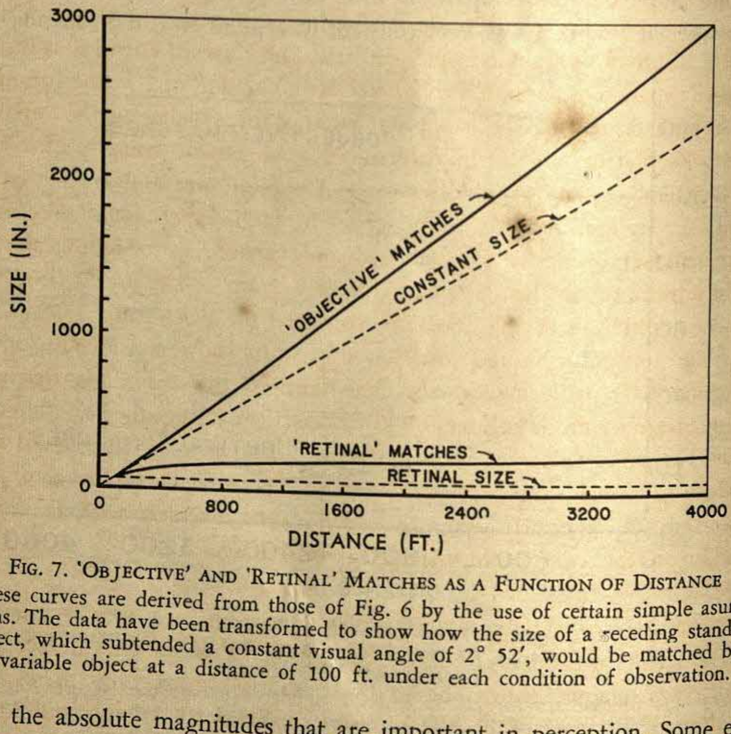


FIG. 7. 'OBJECTIVE' AND 'RETINAL' MATCHES AS A FUNCTION OF DISTANCE
These curves are derived from those of Fig. 6 by the use of certain simple asumptions. The data have been transformed to show how the size of a receding standard object, which subtended a constant visual angle of 2° 52′, would be matched by a variable object at a distance of 100 ft. under each condition of observation.

not the absolute magnitudes that are important in perception. Some evidence for the independence of the size-distance relations from the absolute magnitudes of the stimulus-objects has been given by the present data but the parameter has been varied only within a limited range.

Plotted in this form the functions of Fig. 7 appear to be similar to those of Holway and Boring's Figs. 3-22.[18] The function for the 'objective' attitude is now a straight line rising with distance and exceeding the slope of the line for size constancy. The function for the 'retinal' matches is slightly curvilinear and close to the horizontal slope for perceived size dependent only upon retinal image.

[18] *Ibid.*, 26-34.

At first glance the position of this curve resembles that of the extreme reduction condition of the Holway-Boring study. This is deceiving. Actually, the curve is given by Equation [1] for perceived size in which the value of the parameter A (300 ft.) is in reasonably good agreement with that found to apply to Boring's own binocular observations in the former study. Boring's data were fitted by the same equation with A taken as 243 ft.[19] The higher value of A in the present experiment may be attributed to the many more compelling distance cues given by the airfield in broad daylight as contrasted with the Harvard corridor at night. Neither experiment permitted the direct determination of δ in the formula, but in both cases it has been approximated as the distance of the variable object. Its precise value is immaterial since we assume equality of visual size when object and variable have equal physical size and are placed at identical viewing distance. Some error in both experiments prevented exact matches in size when standard and variable were equidistant. A slight adjustment in the assumed value of δ, from 100 ft. to 90 ft. in the present application of the equation, corrected the discrepancy.

If our binocular 'retinal' matches appear to show considerably more reduction to retinal image size than Boring's, the answer is that our distances are so much greater, 4000 divided by 120 equals 33.3 times as great. For small values of the distance D (relative to A), perceived size is close to actual size, but for very large distances D (relative to A), perceived size is greatly diminished and varies inversely with D, corresponding to the reduction in size of the retinal image.

Discussion

The results show a clear distinction between two attitudes of observation. An S, it appears, may be set by appropriate instructions to respond in alternative ways to the same complex conditions of stimulation. As an object recedes into the distance its size may be judged to get either a little larger or much smaller than its actual size, depending upon what question has been put to S, upon what he means by his judgment.

It is instructive to try to relate these findings to Gibson's two systems, the visual world and the visual field. For Gibson the visual world is the experience of unbounded stable three-dimensional Euclidean space, one in which parallel edges do not converge and in which an object stays constant in size wherever it is moved. An object-directed attitude ought.

[19] Gilinsky, *op. cit.*, 490.

it would seem, to duplicate in perception this natural world of objects. It does not—not exactly.

Matches of 'objective' size do not strictly follow the rule of size constancy. Instead, objects tend under this attitude to *expand* as they move further away. Nor are these data of the present experiment exceptional. Holway and Boring obtained this result from four out of five *S*s but at the time doubted the validity of 'overconstancy' and corrected their data for a presumed space error.[20] Their data for one *S* (*EGB*), even when unadjusted, did not show overcompensation but was typical of the present 'retinal' matches. Presumably the other four *S*s were set to match 'objective' size. At least this inference may be drawn from the nature of the results and from similar data found by Chalmers,[21] who did use explicit 'objective' instructions in a replication of the Holway-Boring experiment. Gibson's own results showed that for every test-object and at every distance except 28 yds., size was overestimated.[22] Smith reported a similar tendency for size to *increase* with distance.[23] In his study *S* was asked to judge the physical size of wooden cubes placed out of doors at 16, 80, and 320 ft. As the physical distance of the standard object increased, *S* required a larger and larger comparison to satisfy his judgment of equality. As in the present experiment, the studies reported by Gibson and by Smith used binocular vision under conditions affording many cues to distance and explicitly set their *S*s for accurate and objective judgments.

There is considerable evidence then that, even when we are set to perceive objects as they really are, we simply do not perceive a visual world that corresponds with rigid accuracy to the physical, tape-measured world. The hypothesis of size-constancy is at best an inexact approximation to a scientific description of the 'visual world.' This discrepancy may surprise those who try to account for perception in terms of its purposiveness. Evolution or the individual's life history, or both, appear to have achieved an organism whose 'best bet' is to overestimate object-size. The amount of overestimation increases with the difference in conditions between the objects being compared, revealing more and more discrepancy between behavior and the basic sensory excitation. Perhaps this tendency toward overadaptation is somehow useful. Still it is difficult to understand how an error in one direction or another could increase

[20] Holway and Boring, *op. cit.*, 34.

[21] E. L. Chalmers, Monocular and binocular cues in the perception of size and distance, this JOURNAL, 65, 1952, 415-423.

[22] Gibson, Motion picture testing, *op. cit.*, 206-207.

[23] W. M. Smith, A methodological study of size-distance perception, *J. Psychol.*, 35, 1953, 143-153.

the organism's chance of survival or be the necessary outcome of past interaction with the environment. The discrepancy, though small, is puzzling, and the evident adequacy of adjustment to the environment should not prevent us from seeking to discover how such adjustment is controlled and limited.

Let us turn to the results obtained with the 'retinal' or 'perspective' attitude. This attitude does indeed seem to give us Gibson's visual field, defined by him as "a picturelike phenomenal experience at a presumed phenomenal distance from the eyes, consisting of perspective size-impressions."[24] Gibson is careful not to identify the visual field with the retinal pattern unmodified, for the values of the visual field depend upon both the dimensions of the retinal image and discrimination of the distance of the perceived object.

In the present experiment the set for matching 'retinal' size never achieved complete reduction to retinal size. The receding triangle was seen to shrink but also to recede. With complete reduction, with all cues to distance eliminated, one would expect perceived size under this attitude to vary with retinal size. Reduction, however, in most of these experiments has been but partial and the cues to distance have been many. Observe the cues in Figs. 1-2. Even in the case of the perception of the full moon's disk, where distance cues would seem to be minimal, the matched comparison disk, a dozen feet away from the *S*, is not reduced as much as is the size of the moon itself on the retina.[25] The perceptual pattern is not the pattern of the retinal image without regard to distance. Instead perceived size appears to depend upon perceived distance in accordance with the mathematical account of visual space developed previously.[26]

The present data clearly show that the experience of the visual field need not depend on the elimination of cues to distance or on long and arduous training but may be achieved readily by attitudinal control. Given many cues to distance, *S* uses such as are basic and compelling, but he may ignore the others and so responds quickly and with assurance.

All of the *S*s reported that they made the 'retinal' or, as they preferred to call them, the 'picture-image' settings with greater ease and confidence than the 'objective' matches, particularly at the larger distances.

The task demanded by the 'objective' instructions appeared to the *S*s

---

[24] Gibson, The visual field and the visual world: A reply to Professor Boring, *Psychol. Rev.*, 59, 1952, 151.

[25] Boring, The moon illusion, *op. cit.*, 59.

[26] Gilinsky, *op. cit.*, 460-482. For an independent derivation of the same formula see G. A. Fry, Visual perception of space, *Amer. J. Optom.*, 27, 1950, 531-553.

as less direct and more interesting. Several of them stated that they proceeded by first setting the variable as if for a 'picture-image' match, then tried to estimate how far away was the standard triangle, and how much bigger it must actually be to look that size and that distance away. One said: 'I might decide that it really is 6 ft. tall. Then I turn to the variable triangle and increase it until it looks 6 ft. tall." Certainly in this instance objective size is being estimated inferentially and is not being immediately perceived.

The *S*s were curious and pressed for knowledge of results following their attempts to judge 'objective' size correctly. The 'retinal' task, on the other hand, seemed a challenge only to their manual dexterity in setting the comparison precisely. Uncertainty with regard to the estimates of 'objective' size was often expressed by such statements as, "That was a sheer guess," or "I'd hate to bet on that."

Unfortunately, no data on intra-individual variability are available to reinforce this reported difference in subjective assurance. The inter-individual variability, in terms of the standard deviations of the average settings, is shown in Table I. These *SD*s appear to indicate closer agreement between *S*s and therefore a more determinate, less highly individual basis for the responses under the 'retinal' attitude, especially at the larger distances. Nevertheless, no sure conclusions about the relative ease or stability of judgments under the two attitudes can be drawn from the data on variability for the reason that the variability of the comparison stimulus was restricted at both extremes.

These results suggest that the classical distinction, drawn by Hering and others, between *perception* (immediate experience) and *estimation* (knowledge or inference) may still be pertinent despite the explanation offered by Gestalt psychology.[27] This school and Gibson regard the phenomena of perception as being directly given when the world of objects is attended to. Object-constancy, they hold, is the natural outcome of ordinary perception and is destroyed only by the artificial reduction of stimulation or by the critical efforts of the trained artist or introspective psychologist.

The present study shows that a sincere attempt to make this supposedly natural phenomenal objective world easily observable can fail, forcing the *S* away from immediacy of judgment into an inference of which he is often quite unsure. The further finding that size-constancy does not hold precisely in the visual world seen under the 'objective' attitude also suggests that the Gestalt conceptions of the phenomenal

[27] Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 289.

world need revision. It required the use of long distances by Gibson, by W. M. Smith, and by the present writer to make the latter discrepancy plain. At close range the well-practiced observations of ordinary existence are quick and sure, creating the false impression of being purely passive affairs. It now appears that we may have a right to call the 'objective' attitude *estimation* and the 'retinal' attitude *perception,* thereby reversing the position of the phenomenologists. Distant objects may be estimated or judged to be larger than their actual size, although there is at present no evidence that conditions can be set up under which they would be 'seen' or 'experienced' as actually expanding when they recede. The 'set' in which size appears to decrease with increasing distance is the stimulus-controlled set for perceived or apparent size in visual space. A corresponding and equally important distinction exists between estimated distance and perceived distance in visual space. The familiar names have the advantage of calling attention to important aspects of the two kinds of data, and the complex problems that remain to be investigated.

The distinction between sense-perception and estimation may be useful in keeping separate the two lines of psychological investigation which Graham has identified as the major directions of current research in perception. One is the study of sensory discriminations and their relations to systematic variations in stimulus conditions. The second and more recent area of interest is the study of discriminative behavior as it is influenced by the attitudes, motives, and past history of the subject. In order to evaluate the results of any experiment it is crucial to determine what kind of data we have. Experiments using similar equipment and apparently similar design may turn out to be solving very different problems and producing entirely discrepant results. Serious confusions may be avoided if a given set of observational data can be identified as belonging to one or another class of discriminations.[28] The particular instructions used, or the fact of their ambiguity, may be as decisive as any other feature of the method of investigation.[29]

Thus an attempted reproduction of the 'physical' value of a stimulus-object may be specified in terms of quantified stimulus-energies; it is nonetheless a stimulus-*rating* and may be shown to possess limitations in common with the verbal estimates or numerical naming responses of the

[28] These considerations may help to clarify the questions raised by W. M. Smith, Gilinsky's theory of visual size and distance, *Psychol. Rev.*, 59, 1952, 239-243.

[29] The discrepant results reported by different investigators of the perception of verticality further illustrate this point. See C. W. Mann and R. O. Boring, The role of instruction in experimental space perception, *J. Exper. Psychol.*, 45, 1953, 44-48.

rating method. The responses are not highly restricted by the instructions and they may show wide individual differences, reflecting differences in attitudes, training, and other variables of *S*'s past activities. Such data may be usefully related to learning, and the individual's habitual pattern of responding. They do not permit the identification of critical stimulus-energies or physiological limits of discriminative capacity. Problems of the latter type need to be investigated by methods which involve a narrowly specified criterion of response. A genuine psychophysics of perception might best concentrate on those data which reveal, as Boring says, "the parametric invariances of the stimulus."[30]

Graham has presented a systematic view of the field of perception which emphasizes the determination of perceptual functions.[31] These functions are obtained by finding how one stimulus variable varies as a function of another in order to produce a constant response effect. The discovery of such functions for invariant perceived size has been clearly recognized by Boring as basic to a science of perception.[32]

With the aid of Equation [1], such functions for size-invariance may be derived readily from the present data for 'retinal' matching. The 'objective' set does not yield data which allow a general description in terms of stimulus-relations.

Consider Fig. 5 again and note the positions of the successive curves. Each curve represents the data obtained for a given standard as its distance varied. One can select a particular value of the dependent variable, say 10 in., and from each curve determine the distance on the abscissa which corresponds to 10 in. on the ordinate. Thus the 10-in. comparison at 100 ft. was matched to the 42-in. standard at 1300 ft., to the 78-in. standard at 3000 ft., and to the intermediate standards at intermediate distances. When we plot the size of each standard against the appropriate distance we obtain a straight line. The extrapolated curve shows how large different objects that look alike (are matched in perception by a 10-in. object) must be if they are from 100 ft. to 4000 ft. away. This relation has been plotted for several arbitrarily selected values of the stimulus variable in Fig. 8.

Fig. 8 thus exhibits a family of perceptual curves derived from Fig. 5 and extrapolated by means of Equation [1] for various response settings. Each derived curve is a rising straight line which describes how the actual size of an object, referrable to a constant response setting, varies as a

[30] Boring, Visual perception as invariance, *op. cit.*, 147.

[31] Graham, Behavior and the psychophysical methods: An analysis of some recent experiments, *Psychol. Rev.*, 59, 1952, 62-70.

[32] Boring, Visual perception as invariance, *op. cit.*, 146 f.

function of the distance of the object from S. The slopes of these lines may be computed directly from Equation [1] for any value of St_c. This plot would yield for size constancy, a family of straight lines of zero slope; and for true visual angle matches, ascending straight lines, but steeper than the corresponding equal size contour for a given perceived size.

This treatment of the data emphasizes that for perceived size to be invariant under the attitude for observing 'retinal' size, neither actual

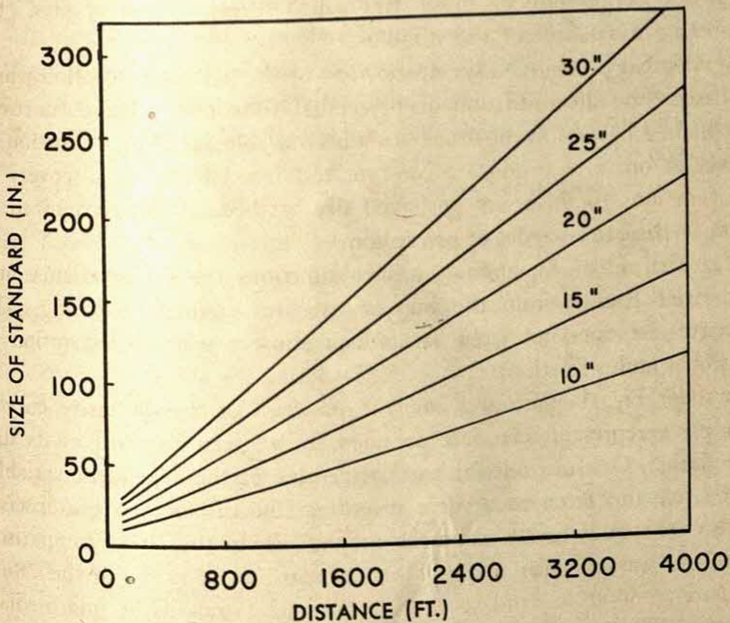


FIG. 8. FUNCTIONS FOR PERCEIVED SIZE INVARIANT UNDER THE 'RETINAL' ATTITUDE

This family of curves is derived from the data of Fig. 6 and Equation [1] for five settings. It shows how the actual size of a test-object must vary with distance to give an invariant response under this attitude.

object-size nor actual retinal size can be invariant; at least, not as long as determinants in addition to retinal size are available. A lawful principle of invariance is disclosed however, which, to paraphrase Boring, tells us how the organism does perceive its own physiological bases; the data out of which it can create, after much evolution and under a sensible, practical, real-life attitude, a sufficiently useful apprehension of the world that it accepts as its reality.[33]

[33] *Ibid.*, 147.

If we designate this manifold a visual field, then surely it is a visual field for partial reduction of cues to distance. Perception under this attitude was always influenced by perceived distance; yet there was no conscious inference that took distance into account, as there was with the 'objective' attitude. Boring suggests that there might be more than one visual world, and also, corresponding to different degrees of reduction, a whole series of visual fields, including the limiting case where reduction is complete and perceived size varies only with retinal size without regard to change in distance.[34] In this way, it seems, the present experiment throws light on Gibson's two visual systems, and suggests even that we may eventually be able to specify, for various observational attitudes, and various degrees of reduction of the cues to distance, an inclusive system of relations to define and explain the phenomena of visual space perception.

SUMMARY

The perception and the estimation of object-size was studied as the distance of a standard object was varied from 100 to 4000 ft., out of doors under conditions affording many cues to distance. The standard stimulus-objects were plane, white isosceles triangles, 42, 54, 66, and 78 in. in base and altitude. Functions relating the adjusted size of a variable triangle, 100 ft. away from *S*, to the distance of each standard triangle were obtained from 32 to 36 *S*s. Each *S* served under two different conditions of instruction which demanded contrasting observational sets; a set for matching 'objective' size and a set for matching 'retinal' size.

(*1*) The data show clearly that instructions were effective in producing two distinct functions relating the settings of the variable triangle to distance.

(*2*) 'Objective' instructions gave matches in size which increased with distance, exceeding size-constancy. The finding that 'objective' judgments overestimate object-size confirms and extends previous studies of size-estimation.

(*3*) 'Retinal' instructions gave matches in size which decreased as distance increased. This relation between perceived size and distance is intermediate between the function for size constancy and the function for matches of retinal image or visual angle. These results are consistent with past findings for apparent or perceived visual size and with a previously developed mathematical formulation of visual space.

[34] Boring, The Gibsonian visual field, *Psychol. Rev.*, 59, 1952, 246-247.

THE EFFECTS OF TERRAIN AND OBSERVATION DISTANCE ON RELATIVE DEPTH DISCRIMINATION

By WARREN H. TEICHNER, JOHN L. KOBRICK, and ROBERT F. WEHRKAMP, Quartermaster Research & Development Command, Natick, Mass.

In an attempt to determine the effects of specific environmental factors on the precision of relative depth discrimination only two studies, one by Hirsch and Weymouth[1] and one by Holway, et al.[2] were found in which this phenomenon was actually studied in an outdoor situation. Although both investigations yielded a considerable amount of information, neither was considered extensive enough to provide good estimates of the effects of viewing distance and of the nature of the terrain over which viewing occurs. The present investigation was performed, therefore, as a further study along this line.

EXPERIMENTAL CONDITIONS

The experimental techniques were based upon the methods of the Howard-Dolman apparatus and in this respect were comparable to both of the studies previously cited. A rectangular board, 77 in. high and 64 in. wide, was mounted on the front bumper of a jeep in a manner which concealed the vehicle from front view. A canvas skirt was hung from the bumper almost to the ground to present a smooth, unbroken target. The jeep then served as the variable stimulus-object. The standard stimulus-object was a board mounted on the ground and comparable to the variable with respect to height above the ground and width. Both targets were painted flat black and presented smooth, unbroken surfaces at 200 ft.

Four soldiers were used as *Ss*. In addition to routine medical requirements, including normal vision (20/20 Snellen, uncorrected), the *Ss* were selected on the basis of a high score on the Verhoeff stereopter.

The administration of the runs was identical for all experimental conditions. *S* stood at one end of the viewing course with the standard target positioned on his right. The variable target was always started at such a distance away from the standard that the two targets were obviously unequally distant. It was then moved toward the standard at the slowest possible speed, its motion being at or below the threshold of perceived motion. The direction of motion of the target was varied in a random manner, but in such a way that an equal number of runs was made in both directions.

---

* Accepted for publication April 25, 1954.

[1] M. J. Hirsch and F. W. Weymouth, Distance discrimination: V. Effect of motion and distance of targets on monocular and binocular distance discrimination, *J. Aviat. Med.*, 18, 1947, 594-600.

[2] A. H. Holway, D. A. Jameson, M. J. Zigler, L. M. Hurvich, A. B. Warren, and E. B. Cook, Factors influencing the magnitude of range-errors in free-space and telescopic vision, Division of Research, Graduate School of Business Administration, Harvard Univ., 1945, 1-314.

At each experimental session the *Ss* were so rotated through trials in a constant order that every *S* rested for three trials following each of his judgments. Between observations the *Ss* rested under a tarpaulin facing away from the targets. Thus, they were unable to watch the variable target being set prior to a run.

At the beginning of every observation the variable was moved slowly toward the standard. When the two targets appeared equidistant, *S* waved a large red flag. At this signal the vehicle was stopped and the error of positioning was measured on the ground with a steel tape-measure. *S* was not allowed to correct his first signal, nor to see the measurement made.

Because of the high illumination present, shadows cast by the targets were not visible at the distances used. Thus, this possible artifact was under control. In addition a lateral separation of 3 min. of visual angle was maintained between the targets under all viewing conditions including movements of the variable target. This value was determined by preliminary experimentation as the smallest separation which could be resolved at a viewing distance of 3000 ft. and still provide 100% detectable target separation for all *Ss*.

Two experiments were performed. They were carried out on the desert near Yuma, Arizona, during July and August, 1953. In the first, measures of error were recorded on a macadamized airstrip at nine different target distances, viz., 200, 475, 750, 930, 1125, 1300, 1500, 2250 and 3000 ft. A photograph of the airstrip with the targets on it is shown in Fig. 1. Only one distance was tested in a single day. Six monocular and six binocular observations were made by each *S* between 8:30 and 11:30 A.M. This entire procedure was repeated in the afternoon of the same day (at the same distance) between 1:30 and 4:30 P.M. Thus, when all *Ss* are considered, two sets of data, each consisting of 24 monocular and 24 binocular observations, were provided at each distance.

The second experiment was performed on three different types of desert terrain. Following a recent geographical classification these were as follows.[3]

(1) Sand plain or desert sand. Surface consists of loose hummocky sand alternating with patches of gravel having a somewhat firmer surface. Color is light tan. Vegetation over the experimental course consists of creosote bush and burro-weed.

(2) Desert pavement. Surface consists of small stones (diameter mostly 2 in. or less) worn smooth on exposed face and fitted closely together forming a continuous 'paved' surface except where broken occasionally by vehicle tracks. Surface patina or 'desert varnish' gives the stones a shiny, dark brown to nearly black color. Some scattered loose stones are present on the surface. No vegetation grows on the surface, but small drainage channels crossing it have sparse growths of palo verde and iron-wood trees.

(3) Silt. This area consists of alluvial deposits bordering the Gila River. The surface is tan in color and has a thin crust when dry which readily becomes loose and powdery if disturbed. Vegetation is almost lacking at the site used, there being only a few clumps of burro-weed and saltbush present on the entire course.

The actual sites on each desert terrain were chosen with the one restriction that there be nothing present to obstruct a complete view of the targets at any time. Measures on these terrains were taken at observation distances of 200, 750, and 1500 ft. All other details of the second experiment, *e.g.* monocular-binocular viewings, morning-afternoon measurements, etc., were identical to those of the first experiment.

[3] Handbook of Yuma Environment, EPB Report No. 200, OQMG, 1953.

Although two experiments were carried out, actually they were conducted as one. That is, measurements were taken in an irregular fashion with respect to the terrain, including the airstrip, and the distance used on a particular day. Thus, special weather effects could be expected to be non-systematic with respect to the experimental conditions. As it happened the weather was very stable.

RESULTS

The results were analyzed first in terms of the constant error (*CE*) of the equality settings. In each experiment the statistical sign-test was applied in two ways:[4] (1) to test the null hypothesis that 'near' settings—those on *S*'s side of the standard—did not differ in directional sign from the 'far' settings—those beyond the standard—for the various conditions

FIG. 1. AIRSTRIP: TARGETS AT 750 FT.

of the experiment; and (2) that regardless of directional sign, 'near' settings did not differ in magnitude from the 'far' settings. The first null hypothesis was rejected in both experiments with a probability of less than 0.01. From this it could be stated that 'near' and 'far' settings did differ in directional sign, *i.e.* that there was a different kind of *CE* associated with

[4] J. W. Dixon and F. J. Massey, *Introduction to Statistical Analysis*, 1951, 247-254.

each. The second null hypothesis was accepted for both experiments since the differences in magnitude were significant at a probability level of greater than 0.25. Thus, it was possible to conclude that measures taken over both the 'near' and 'far' settings combined would effectively cancel out all systematic biases of the settings, *i.e.* that the *CE*s obtained from the combined near and far settings were random deviations from zero.

Having demonstrated the absence of uncontrolled, systematic effects in the investigation, the data were analyzed in terms of the standard deviation (*SD*) of the combined near and far settings. This statistic represents the distribution of approximately two-thirds of the scores about the obtained *CE* for a given experimental condition and by inference is the

TABLE I

ANALYSIS OF VARIANCE OF AIRSTRIP DATA

Source	*SS*	*df*	*MS*
Distance	31.4104	8	3.9263*
AM–PM	0.0262	1	0.0262
Mon.–Bin.	0.4246	1	0.4246†
Dist. × AM–PM	1.4320	8	0.1790
Dist. × Mon.–Bin.	0.0913	8	0.0114
Mon.–Bin. × AM–PM	0.0251	1	0.0251
Residual error	0.6965	8	0.0871
Total	34.1061	35	
Av. error	2.1449	25	0.0898

* Significant at 1-% level. † Significant at 5-% level.

best estimate of the population of the settings around the population mean of zero. In the present context it will be used as the estimate of the threshold of relative distance discrimination in linear measure.

Since the *SD*s were based on small samples, normality could not be assumed for the application of variance analysis procedures. For purposes of statistical test, therefore, the *SD*s were transformed to normalized ranks with the use of tables provided by Fisher and Yates and analysis of variance was performed upon the rank measures.[5] A summary of this analysis for the data obtained from the airstrip is presented in Table I.

It may be determined from Table I that none of the interaction variances is significant when tested over the residual error. On the asumption that the interaction variances are really random effects, a better estimate of the experimental error may be obtained from a pooled average of all interactions, including the residual error. Such an error-term was calculated and is shown as the average error of Table I. With this as our estimate of

[5] R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, 1948, 66.

error, Table I indicates that variations in distance and kind of viewing, whether monocular or binocular, were the only significant main effects on the airstrip.

Fig. 2 shows the linear threshold (*SD*) as a function of distance for measures taken along the airstrip. Every *SD* is based on 48 measures, the combined morning and afternoon observations. The thresholds are shown separately for monocular and binocular viewing. It will be noted that the 1500- and 2250-ft. values of this figure were treated as different from the general trend and connected separately. This was done for two reasons: (1) all distance locations along the airstrip were level except these two; these two distances lay on a slight rise or slope of approximately one to two degrees, a fact which is noticeable in Fig. 1; and (2) a second study performed in the laboratory confirmed the hypothesis that a rise from the level of as little as was present might be expected to produce a marked effect on the threshold.[6] For these reasons further analysis will treat these two points independently of the rest of the data.

Fig. 2 shows that the linear threshold increased with distance over the range tested from approximately 5 to 100 ft. The smooth lines exhibited in each part of the figure are exponential curves fitted to the data. Equations based on the theoretical expectation, $f(D^2)$, were found *not* to fit the data. The functions shown in Fig. 2 are: Monocular Threshold $= 0.005\ D^{1.24}$; Binocular Threshhold $=\ 0.0006\ D^{1.517}$. These functions may be seen to provide fairly good fits to the data of Fig. 2.

Fig. 3 presents the same data as shown in Fig. 2, but in addition allows for a comparison of the relative efficiency of monocular and binocular viewing. From the upper half of this figure, where the empirical values are shown, it can be seen that binocular viewing was slightly superior to monocular, up to at least 1300 ft. The data also show that binocular superiority was lost somewhere before 3000 ft. The only unexpected inversion shown in the upper half of this figure is at the shortest viewing distance, 200 ft., where the monocular measure was actually slightly less than the binocular one. The lower half of this figure presents a monocular-binocular comparison in terms of the fitted curves of Fig. 2. Here it may be seen that binocular viewing appears to be superior until about 1900 ft. Also these curves indicate that the binocular threshold should be lower than the monocular at 200 ft., an expectation which, as noted, is not in accord with the obtained values shown in the upper half of the figure.

[6] E. R. Dusek, W. H. Teichner, and J. L. Kobrick, The effects of the angular relationships between the observer and the stimulus-surround on relative distance discrimination (in prep.).

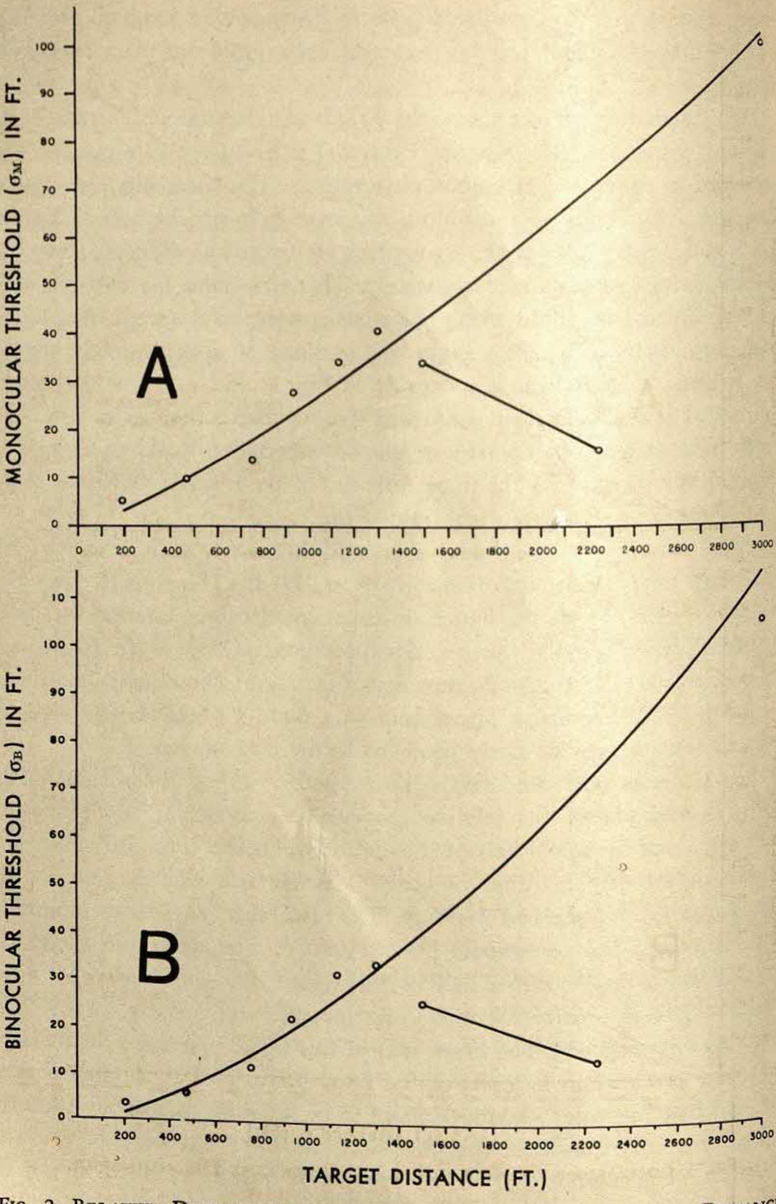


FIG. 2. RELATIVE DISCRIMINATION AS A FUNCTION OF OBSERVATIONAL DISTANCE
(A) Monocular; (B) Binocular

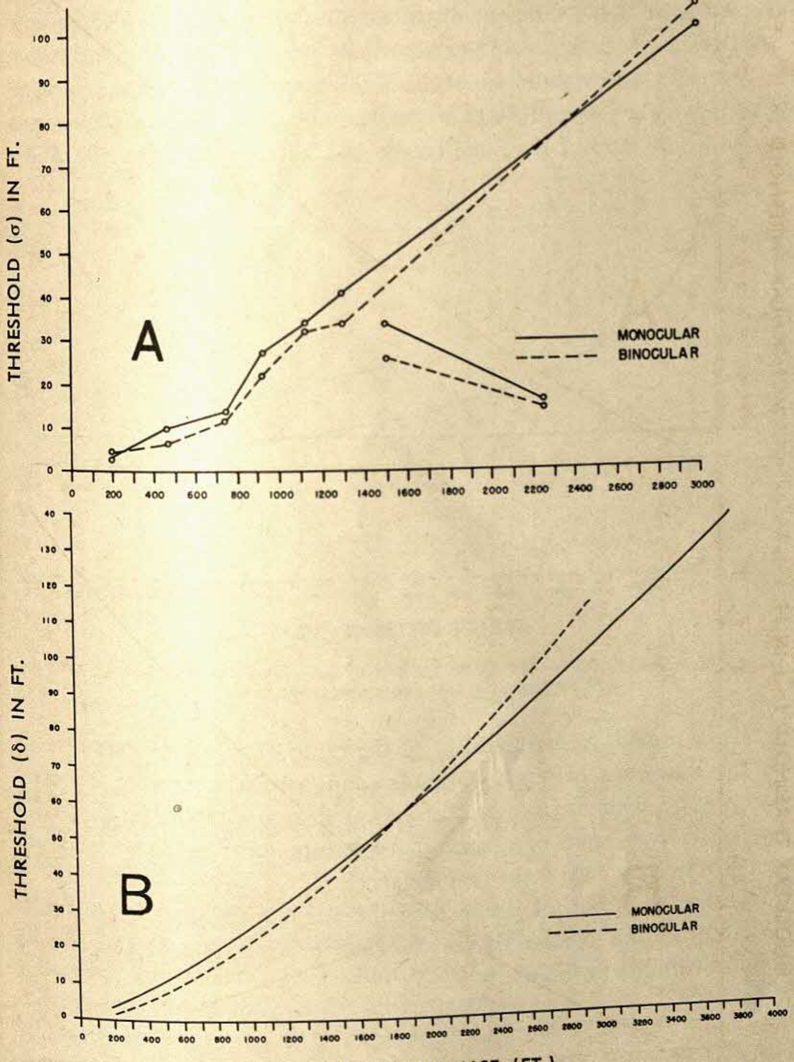


FIG. 3. COMPARISON OF MONOCULAR AND BINOCULAR DISCRIMINATIONS
(A) Empirical; (B) Fitted curves

A further test of the adequacy of the fitted lines may be provided by comparing the ratio of the monocular to the binocular function with the obtained ratio at the different distances. This comparison is made in Fig. 4. Where the ratio equals 1.00, shown by the horizontal line of Fig. 4, the two types of threshold are equal. Values above 1.00 indicate binocular superiority. The smooth line drawn through the empirical values was derived as the ratio of the fitted trends and, except for the 200 ft. values,

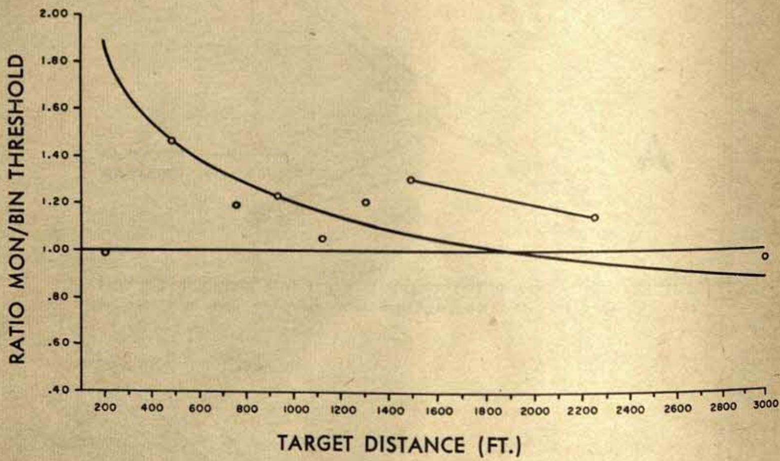


FIG. 4. RATIO OF MONOCULAR TO BINOCULAR DISCRIMINATIONS AS A FUNCTION OF OBSERVATIONAL DISTANCE

appears to provide a reasonable fit of the ratios. This figure shows clearly that the difference between monocular and binocular viewing decreased systematically with distance and disappeared around 1900 ft. Further, it may be observed from this figure that the superiority of binocular vision was not very great even at the near distances.

To assess the effect of distance on the angular threshold of binocular disparity, the *SD*s obtained from the airstrip were converted to angular measures with the formula:

$$\eta_+ = [(IP)\ (SD)\ K]/D^2,$$

where: IP = interpupillary distance; SD = linear threshold; $K = 206265$; D = observation distance; and η_+ = Threshold of binocular disparity in seconds of arc. Empirical values of η_+ were obtained by using the actual *SD*s in the above formula[7]; computed values were found using 0.0006

[7] C. H. Graham, Visual perception, In S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 888, presents a recent discussion of this equation.

$D^{1.517}$ values. This formula was used to compare the obtained angular thresholds with those which should be expected on the basis of the curve depicted in Fig. 2-B. The data are shown in Fig. 5. It will be noted that this theoretical curve provides a reasonably good description of the empirical values although it is low at the 200 ft. value. This suggests that the obtained value for the 200 ft. distance was in error by being too high, a suggestion

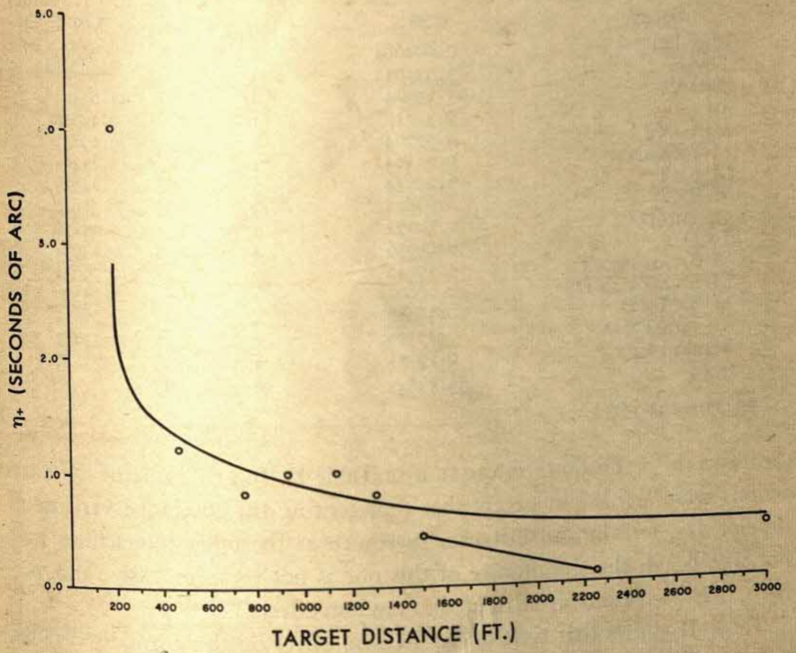


FIG. 5. THRESHOLD OF BINOCULAR DISPARITY AS A FUNCTION OF OBSERVATIONAL DISTANCE

which is consistent both with this figure and previous ones. Fig. 5 shows that the angular threshold decreased with increases in distance falling rapidly at first from about three or four sec. of arc and then less rapidly to approximately 0.50 sec. of arc. Also this figure shows that the values obtained at 1500 and at 2250 ft. were extremely small. These values are 0.47 and 0.11 sec. of arc, respectively.

Table II presents the summary of an analysis of variance of the normalized ranks of the *SD*s obtained from a comparison of the three desert terrains. As shown in this table, the effects of the terrain and of the observation distance were highly significant. Fig. 6 shows these effects separately for monocular and binocular viewing and, in addition, allows for a

comparison of the terrain thresholds with the curves fitted to the airstrip data. As may be observed, the differences between terrains, although significant, are small and not completely consistent for the two kinds of viewing. This lack of consistency is not supported, however, by a significant Terrain × Distance interaction in Table II. In fact, none of the interactions were significant except that representing the combined variance of time-of-day, terrain and distance. Inasmuch as the other interactions were not significant, the significance of this one is not very impressive and, perhaps, may be reasonably attributed to random effects.

TABLE II

ANALYSIS OF VARIANCE OF DESERT TERRAIN DATA

Source	*SS*	*df*	*MS*
Mon–Bin	0.004694	1	0.04694
Terrain	3.55562	2	1.77781*
Distance	25.25024	2	12.62521*
AM–PM	0.34810	1	0.34810
AM–PM×T	0.35704	2	0.17852
M–B×AM–PM	0.22404	1	0.22404
M–B×T	0.49654	2	0.24827
M–B×D	0.65824	2	0.32912
AM–PM×D	0.10271	2	0.05136
T×D	0.99736	4	0.24934
M–B×AM–PM×T	0.19942	2	0.09971
M–B×AM–PM×D	0.06601	2	0.03300
M–B×T×D	0.36948	4	0.09237
AM–PM×T×D	0.55115	4	1.37787*
Residual error	0.29353	4	0.07338
Total	33.51642	35	

* Significant at 1-% level.

Table II shows that time-of-day had no effect on the normalized ranks, a fact which is in agreement with the airstrip data. Every *SD* of Fig. 6 is, therefore, based on 48 observations, *i.e.* the combined morning-afternoon data. Table II also indicates that there was no difference between monocular and binocular observations, a result which does not agree with the findings on the airstrip. In actuality, the binocular thresholds were slightly lower than the monocular ones on the Desert Plain and on the Desert Silt. On the Desert Pavement, however, there was no difference between the two at 200 and 750 ft. and at 1500 ft. the binocular value was slightly higher.

DISCUSSION

Although the present study has not demonstrated differences between morning and afternoon observations, it cannot be concluded that such

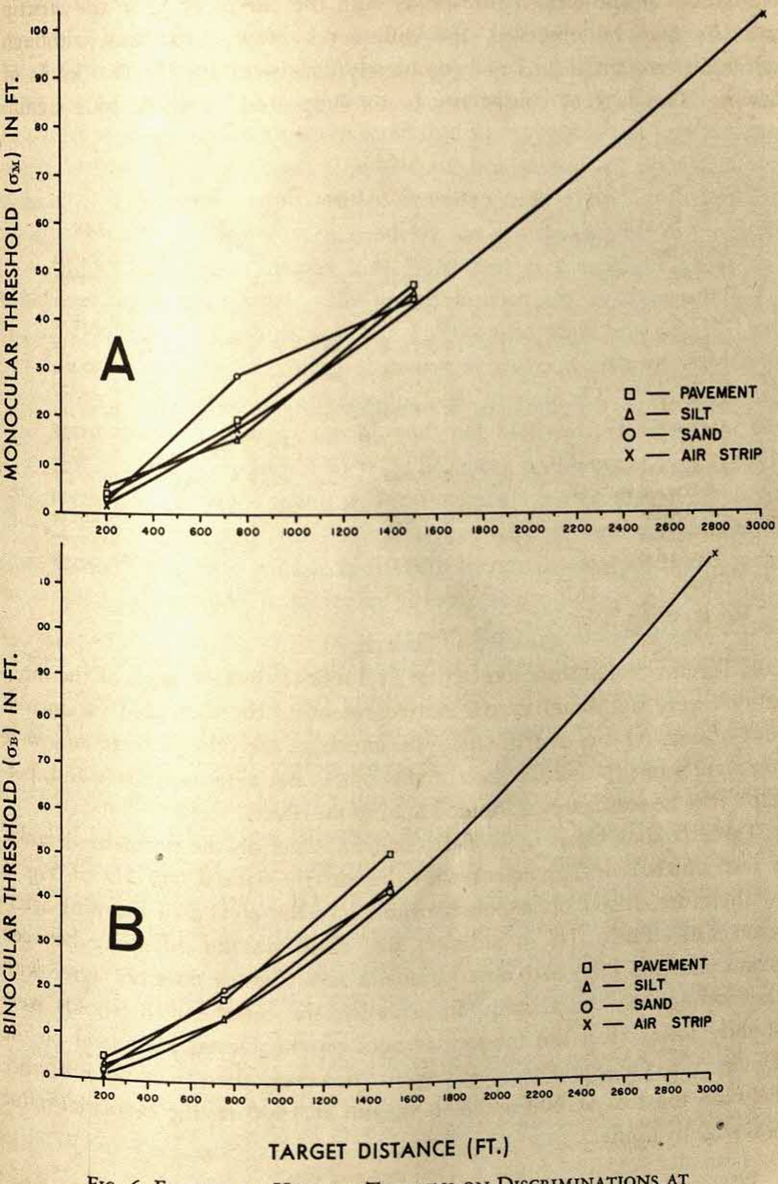


FIG. 6. EFFECTS OF VARIOUS TERRAINS ON DISCRIMINATIONS AT DIFFERENT OBSERVATIONAL DISTANCES
(A) Monocular; (B) Binocular

variables as illumination and atmospheric boil have no effect. Other investigations have shown that illumination is of great importance in relative distance discrimination.[8] Since no measure of atmospheric boil was available in the present investigation, it cannot be stated whether or not there were changes in the amount of boil from morning to afternoon or from one type of terrain to another and, therefore, it cannot be said whether or not an effect should have been expected. Although atmospheric boil has been measured in the past, it has not yet been related to depth discrimination.[9] The possibility that it is so related must remain an open question.

On the basis of geometrical theory Cibis has deduced the hypothesis that, "if size and slant of a surface are kept constant, both binocular and monocular parallax decrease in proportion with the square of the observation distance."[10] The present data support the expectation of a consequent rise in the linear threshold but they do not support the expectation of a rise as rapid as would be produced by $f(D^2)$. The studies of Holway et al. are in agreement with the present one that under commonplace viewing the thresholds rise less rapidly with distance than theory predicts.[11] On the other hand, the present results differ in that they suggest a distance function, $f(D^{1.5})$ for binocular viewing rather than $f(D^j)$, $j \leq 1.00$, as reported by Holway et al.[12]

Inasmuch as the binocular threshold rises more rapidly with distance than the monocular, it is not surprising that the two curves cross. The point at which they cross, however, is far short of the 4000 ft. predicted by Hirsch and Weymouth,[13] and considerably longer than the 900 ft. predicted by Cibis.[14] Graham, on the other hand, in making a rough guess purely for purposes of illumination estimated about 1500 ft. as the limit of stereoscopic vision.[15] It is interesting that this guessed-at limit is so close to our theoretical limit of 1900 ft. Of course, the actual limit of discrimination based solely on transverse disparity is dependent to a large degree on individual differences in stereoscopic acuity.

Despite the relatively long range over which binocular viewing was

---

[8] J. G. Beebe-Center, Leonard Carmichael, and L. C. Mead, Daylight training of pilots for night flying, *Aeronaut. Eng. Rev.*, 3, 1944, 1-10.

[9] L. A. Riggs, C. G. Mueller, C. H. Graham, and F. A. Mote, Photographic measurements of atmospheric boil, *J. Opt. Soc. Amer.*, 37, 1947, 415-420.

[10] P. A. Cibis, Problems of depth perception in monocular and binocular flying, *J. Aviat. Med.*, 23, 1952, 619.

[11] Holway et al., *op. cit.*, 1-7.

[12] *Idem.*

[13] Hirsch and Weymouth, *op. cit.*, 596.

[14] Cibis, *op. cit.*, 619.

[15] Graham, *op. cit.*, 889.

superior to monocular viewing on the airstrip, the ratio of superiority was not very great. Hirsch and Weymouth obtained ratios of this sort at 200, 400 and 600 ft. of 2.81, 2.04 and 1.77, respectively.[16] Their ratios roughly exhibit the same trend as ours, but are somewhat higher. The present ratios derived from the fitted curves for the same distances declined from 1.88 to 1.38. The differences in the size of the ratios obtained from the two experiments are small and possibly are the result of differences in method. For example, Hirsch and Weymouth made their observations looking up a slope whereas ours were made over a level surface.[17] It is possible that this difference in slope is largely responsible for the discrepancy between the two sets of data. The rise of the airstrip at 1500 and 2250 ft. in the present study appears to be an independent variable and one which was responsible for marked changes in both types of threshold.

Beebe-Center, Carmichael, and Mead in a study of the effects of outdoor illumination on depth perception reported that although the linear binocular threshold was larger at a viewing distance of 200 ft. than it was at 100 ft., the angular threshold was smaller at the larger distance.[18] The implication of this finding is, of course, that stereoscopic acuity *improves* with distance. This implication is consistent with the finding of Holway et al. that the percentage of error in setting decreases with increases in observation distances and is clearly confirmed by the present investigation.[19] Our data suggest an angular disparity of 0.64 sec. of arc as the lower threshold at the limiting distance, *i.e.* the distance at which binocular superiority is lost. These data are confounded, however, by the presence of monocular cues and this is undoubtedly much too small a value for pure stereopsis. In fact, reference to the angular disparities obtained at 1500 and at 2250 ft. on the airstrip show that disparities as small as 0.11 sec. of arc are obtainable in the presence of environmental factors. These appear to be incredibly small values until one compares them with parallactic angles of 0.03 and 0.05 sec. of arc reported by Holway et al.[20]

The differences obtained among the various terrains although significant were small and not consistent from terrain to terrain or from distance to distance. The most marked effect of the terrain appears to be the lowered thresholds produced by the change in slope on the airstrip. That

[16] *Ibid.*, 594.
[17] *Ibid.*, 595-596.
[18] Beebe-Center, Carmichael, and Mead, *op. cit.*, 1-10.
[19] Holway et al., *op. cit.*, 1-314.
[20] *Idem.*

such changes in slope may markedly affect the threshold was confirmed in an independent study.[21] This effect suggests that the angle of view formed by the eye at the terrain immediately about the targets (base-surround of the targets) is an important variable. It is likely that all angular relationship involved in forming the total angle of view are of importance, *e.g.* eye-level height over the ground, slope of the base-surround with respect to the level, target tilt, etc. and that they should receive greater emphasis than heretofore.

One explanation of the results likely to be common is that *S*s discriminated in terms of differences in retinal size of the target. Such differences would become larger as the distance between the targets was increased. It could be hypothesized, moreover, that not only is difference in size a factor, but that actual size, itself, has an effect on the sensitivity to difference in size. Woodburne, however, reported that, for binocular vision, changes in visual angle resulting from movement of the variable stimulus in the Howard-Dolman test are at best an accessory factor of minor importance.[22] Graham, Baker, Hecht, and Lloyd presented clear-cut evidence showing that differences in visual angle produced by movement of the variable of two target needles have no effect.[23] Vernon's data are also consistent in this regard.[24] On the other hand, Holway et al. interpreted their data as indicating that size is a primary factor.[25]

It may be that size difference is not the major cue, but rather that something partly dependent upon it is. More specifically, in our situation, it could be hypothesized that monocular discriminations were made in terms of a coincidence matching or vernier alignment of some aspect or aspects of the targets, *e.g.* the top edges. Such an hypothesis assumes that difference in retinal position rather than retinal size constitutes the primary cue. The *S*s in the experiment reported by Graham, Baker, Hecht, and Lloyd actually reported something like this as the major basis for their discrimination.[26] In that experiment, furthermore, it was shown that judgments so based were independent of the relative sizes of the stimuli. This hypothesis would explain the exceptionally low thresholds obtained at the 1500- and 2250-ft. distances on the airstrip for at these points

[21] Dusek, Teichner, and Kobrick, *op. cit.*, (in prep.).

[22] L. S. Woodburne, The effect of a constant visual angle upon the binocular discrimination of depth, this JOURNAL, 46, 1934, 273-286.

[23] C. H. Graham, K. E. Baker, M. Hecht, and V. V. Lloyd, Factors influencing thresholds for monocular movement parallax, *J. Exper. Psychol.*, 38, 1948, 205-223.

[24] M. D. Vernon, The perception of distance, *Brit. J. Psychol.*, 28, 1937, 1-11.

[25] Holway et al, *op. cit.*, 1-7.

[26] Graham, Baker, Hecht, and Lloyd, *op. cit.*, 212-214.

differences in retinal position would be expected to be exaggerated as compared to targets on level ground. It would also account for the small and inconsistent differences among the terrains for although these surfaces were approximately the same in average contour, slight changes in contour would not occur at the same places along the courses.

It should be observed that the present hypothesis essentially reduces monocular relative depth discrimination to a case of vernier acuity, *i.e.* vernier acuity defined as the minimal resolvable misalignment of two lines. Accordingly binocular discrimination should depend on both vernier acuity and transverse disparity. This hypothesis requires further test.

According to Gibson the density gradients of the targets and of the terrains should have an influential effect on the type of measurement being reported.[27] Since the targets presented highly homogeneous surfaces to *S*, it seems unlikely that there were differences in target density gradient sufficient to provide cues to relative distance. This conclusion is based upon a recent study by Gibson and Robinson in which identically textured surfaces viewed at different distances failed to produce an impression of relative depth.[28] Further, although there is no doubt that there were differences in density gradient over the various terrains, it does not seem as though these differences should have any effect *up to the target area.* It is only between targets that differences in terrain density would appear to be of importance. It does not seem, however, as though such differences, although present, were of much consequence.

The terrains involved also presented differences in the number and kind of irrelevant stimuli lying between the targets. Gogel has suggested that since the presence of irrelevant stimuli produces a more complex set of retinal disparities than is the case when such objects are absent, equality setting of two targets might be more difficult.[29] Gogel has further suggested that the increase in error to be expected depends on the size of the irrelevant object relative to the sizes of the targets and on its relative proximity to one or the other target.[30] Although the present study did not systematically provide data on this matter, some variations in the

---

[27] J. J. Gibson, *The Perception of the Visual World,* 1950, 1-235.

[28] J. J. Gibson and J. S. Robinson, The effectiveness of an abrupt change in texture-density as a stimulus for the impression of depth-at-an-edge: Human Resources Research Center, Air Training Command, Research Note P & MS, 52-2, 1952.

[29] W. C. Gogel, The perception of the relative depth position of objects as a function of other objects in the field of view. Army Med. Res. Lab., Fort Knox, Ky., Report No. 107, 1953.

[30] Gogel, *loc. cit.*

kind and amount of irrelevant intervening stimuli were present. These variations were produced by differences in the characteristics of the various terrains, specifically differences in the kind and amount of vegetation present on each course. It might be that these were important factors in producing the differences which were observed among the terrains. However, considering the relatively large differences in vegetation between some of the terrains, the differences in discrimination obtained appear quite small.

Summary and Conclusions

In an experimental situation modeled after the Howard-Dolman apparatus *S*s made equality judgments of the relative spatial positioning of two large targets when the targets were at different observation distances and on different types of terrain surfaces.

Analysis of the results showed that the linear threshold of equality increased exponentially with viewing distance for both monocular and binocular viewing. Prediction of the linear threshold based on the exponential function which expresses stereoscopic acuity was not found to be accurate for commonplace viewing conditions. A modification of the formula for outdoor viewing was suggested.

The linear threshold was also found to vary slightly among the terrains used. Further, the results indicated that binocular viewing is generally superior to monocular viewing, although the differences are small and may not always be present. When binocular superiority is present, however, it may be expected to be lost by approximately 1900 ft.

The data also showed that the threshold binocular disparity is not a constant, as might be assumed on the basis of pure stereopsis, but rather that it decreases systematically with increases in distance. It was found that in the presence of environmental factors, extremely small threshold disparities may be obtained.

SYMMETRY, INFORMATION, AND MEMORY FOR PATTERNS

By Fred Attneave, Lackland Air Force Base

It has been generally held by Gestalt psychologists that 'good' figures are remembered more accurately than 'poor' ones. The trace-system resulting from a well-organized field is said to be more stable than that arising from a chaotic one, and hence less subject to effects of interference.[1]

In a recent article the author demonstrated that various Gestalt-factors including symmetry, good continuation, and other forms of regularity may all be considered to constitute redundancy in visual stimulation, and quantified accordingly within a framework of information-theory.[2] Hochberg and McAlister independently arrived at a similar point of view from a consideration of conditions involved in the perception of depth.[3] This redefinition of 'good figure' in terms of redundancy, or interdependencies among parts, raises some new and interesting questions concerning the memory for form. Thus we may ask: Are regular figures better remembered than irregular ones simply because they contain less information to be remembered, or does their superiority persist even when information is held constant? In other words, which is remembered more accurately; a large, well-organized figure, or a small, poorly-organized figure containing the same amount of information?

An experimental investigation of this problem is reported here. Two initial decisions were necessary: (1) a choice of the kind of redundancy, or regularity, to be varied; and (2) a decision as to the variety of memory that would be measured. *Symmetry* was chosen for study, in preference to other forms of redundancy, because its quantification in physical terms is particularly easy and obvious. Three separate experiments were performed to investigate the rôle of symmetry in three different kinds of memory: immediate reproduction, delayed reproduction, and identification. The same basic method of varying stimulus-materials was employed in all three

* Accepted for publication April 12, 1954. This research was carried out under the Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas, in support of Project 7706, Task 77117. The assistance of Mr. D. F. DeVoto in the collection and statistical analysis of the data reported here is gratefully acknowledged.

1 Kurt Koffka, *Principles of Gestalt psychology*, 1935, 507-508.

2 Fred Attneave, Some informational aspects of visual perception, *Psychol. Rev.*, 61, 1954, 183-193.

3 J. Hochberg and E. McAlister, A quantitative approach to figural "goodness," *J. Exper. Psychol.*, 46, 1953, 361-364.

experiments to determine the extent to which effects of symmetry are associated with the reduction of information.

Experiment I: Immediate Reproduction

Subjects. A total of 149 airmen (basic trainees) at Lackland Air Force Base served as *Ss*. They were divided among 5 experimental groups varying in number from 28 to 31.

Materials. The essential experimental variables were provided by construction of different stimulus-materials for the 5 groups. In every case the materials consisted of patterns of dots within rectangular matrices. For the first set of patterns a matrix of 12 cells (3 cells in width and 4 in height) was used. The occurrence or non-occurrence of a dot in any cell of a particular pattern was determined independently, and with a probability of 0.5, by means of a table of random numbers. Sixty patterns,

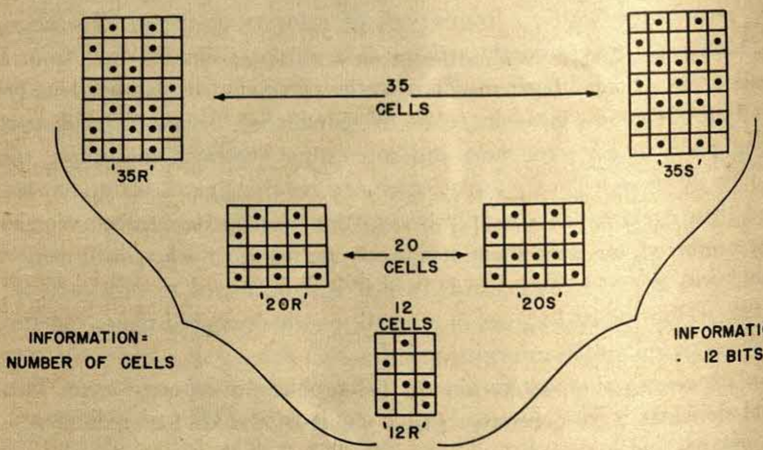


Fig. 1. Examples of the Five Classes of Patterns

The symmetrical patterns 20 S and 35 S are matched with random patterns 20 R and 35 R. With respect to informational content, they are matched with each other and with 12 R: *i.e.* 12 R contains all the information necessary for the construction of either 20 S or 35 S

statistically independent of one another, were made up in this way. The total number of dots per pattern necessarily varied in accordance with the binomial distribution, about a theoretical mean of 6. The label '12 R' will be used to designate this class of 12-cell random patterns. A pattern representative of the class is shown at the bottom of Fig. 1, to which the reader should refer throughout the present discussion.

A second set of 60 patterns was constructed within a matrix of 20 cells (5 cells in width and 4 in height). Over the first 3 columns, these patterns were identical with those of set 12 R, *i.e.* identical cell by cell, not merely chosen by the same method; the occurrence of dots in the remaining 8 cells was determined by the further use of a table of random numbers. This class will be called '20 R.' A third class, '35 R,'

was constructed within a matrix 5 cells wide and 7 high. These patterns were identical with those of Class 12 R in the 12 cells occupying the upper left corner of the matrix, and with those of Class 20 R over the first four rows; the remaining 15 cells of every pattern were again dotted randomly.

Patterns of the remaining two classes were symmetrical, and accordingly redundant. In Class 20 S, the first 3 columns were identical with the patterns of 12 R (and with the first 3 columns of 20 R). The remaining 2 columns were not determined randomly, however: column 4 was a repetition of column 2; and column 5 a repetition of column 1. Thus the patterns were bilaterally symmetrical about an axis running through the third or middle column. The patterns of the remaining class, 35 S, were symmetrical vertically as well as horizontally, with axes of symmetry passing through the third column and the fourth row. As in the case of 20 S, they were constructed by merely 'reflecting' the patterns of 12 R.

Within any one of the random classes (12 R, 20 R, or 35 R) the number of bits of information per pattern was precisely equal to the number of cells, since the state of every cell was determined by a separate dichotomous decision, taken from a table of random numbers. Within the symmetrical classes, however, every pattern contained just 12 bits of information, since it was completely determined by the same 12 random decisions which went into the construction of some pattern in 12 R. Thus, a comparison of memory for the symmetrical patterns with memory for the patterns of Class 12 R should indicate the effect of symmetry with information held constant; whereas a comparison between the memory for symmetrical and random patterns containing the same number of cells should indicate the effect of symmetry when it involves a concomitant reduction in information.

The reader may wonder why symmetrical patterns were constructed about axes of symmetry passing through a row and a column, instead of by repeating all 12 cells of the basic patterns symmetrically. The latter alternative was rejected chiefly because it was considered that a random pattern of 48 cells, which would be required to match the symmetrical pattern produced by doubling the original pattern twice, would exceed the bounds of reasonable difficulty in some of the tasks to be studied. The size and proportions of the smallest matrix, as well as the axes of symmetry, were chosen with the additional aim of keeping all three matrices as nearly alike in length-width ratio, or compactness, as other considerations would permit.

The 5 sets of 60 patterns described above were all redrawn in a 'negative' version in which the state of every cell was reversed, *i.e.* every empty cell in the original patterns became a dotted cell in the 'negative' version, and vice versa. It is obvious that the 'negatives' had the same statistical characteristics as the 'positives,' because of the p of 0.5 originally employed in the dotting of cells. These parallel versions of the patterns made possible a counterbalancing technique which will presently be described.

Apparatus. Patterns were presented to the *Ss* in small groups of 7–8 by means of an apparatus designed by French.[4] The only part of this apparatus which is seen by *S* is a large translucent screen on which points of light, separated vertically and horizontally by multiples of 1 in., may be presented. Patterns were seen by *S* within a clearly marked grid of either 12, 20, or 35 squares, as in Fig. 1. The grids were

[4] R. S. French, The discrimination of dot patterns as a function of number and average separation of dots, *J. Exper. Psychol.*, 46, 1953, 1-9.

made by ruling the desired number of 1-in. squares on small sheets of plexiglass with a blue wax pencil. These sheets could be interchangeably attached to the screen of the apparatus. The grid appropriate to the class of patterns being displayed remained in full view of *S*s throughout an experimental session. The duration of exposure of individual patterns was controlled by means of a Hunter interval-timer.

Procedure. *S*s were assigned unsystematically to five approximately equal groups corresponding to the five classes of patterns. The groups were all treated alike except that every group was given a different class of patterns; accordingly the same designations will be used for groups as for classes of patterns (*i.e.* 12 R, 20 S, etc.). Every group was divided into 4 subgroups of 7 or 8 *S*s. Within any group, the patterns presented to Subgroups C and D were the negatives of those presented to A and B. The order in which patterns were presented to Subgroups A and C was reversed for B and D. A separate experimental session was therefore required for every Subgroup.

At the beginning of a session, the *S*s were seated about 15 ft. in front of the screen, and were given small booklets in which 60 blank matrices of appropriate size were mimeographed. They were told to observe carefully the patterns which were to appear in the grid on the screen, and to try to reproduce, on the mimeographed form, an accurate copy of every pattern immediately after its presentation. A spoken 'ready' signal preceded the exposure of every pattern; the duration of exposure was 1 sec., and the interval between exposures, during which *S* attempted to reproduce the pattern and *E* manipulated the necessary switches for presenting the next pattern, was 30 sec. The first 10 patterns of the series were considered to constitute 'practice' trials, and were so represented to *S*s. It was considered to be of some importance that *S*s—particularly those reproducing symmetrical patterns—establish a set or expectation appropriate to the class-characteristics of the patterns they were given before being presented with patterns on which their performance was to be measured.

It obviously would not have been meaningful to compare the five groups with respect to total errors, since the number of opportunities for error varied widely between groups. This difficulty was resolved by scoring every pattern over only 12 cells—the 12 in which patterns of all classes were identical (see Fig. 1). All other errors were completely ignored: the results which follow were obtained by using the same scoring keys for all groups, and in every case errors have the same possible maximum of 12 per pattern, and the same chance-level of 6 per pattern.

Results. An analysis of variance showed highly significant differences among the five groups. The very high *F*-ratio of 83.22 ($df = 148/4$) was obtained for over-all differences, and an even higher *F* of 85.85 ($df = 86/2$) for differences between groups 12 R, 20 S, and 35 S. A *t* of 2.03, significant at the 5-% level, was obtained between Groups 20 R and 20 S, and a *t* of 2.76, significant at the 1-% level, between Groups 35 R and 35 S.[5]

[5] Both here and in subsequent analyses, the error term (the 'within' s^2) was obtained by summing squares over all the members of each group without regard for subgroups; hence it contains a possible 'between-subgroups' component. Such a pro-

Mean error scores of the various groups are shown in Fig. 2. The results may be considered to describe two functions: one of these contains the 3 random groups, and shows the effect of increasing number of cells and information concomitantly; the other, containing Group 12 R and the 2 symmetrical groups, shows the effect of increasing number of cells without increasing information. From the latter it is evident that symmetrical patterns are considerably more difficult to apprehend than random ones

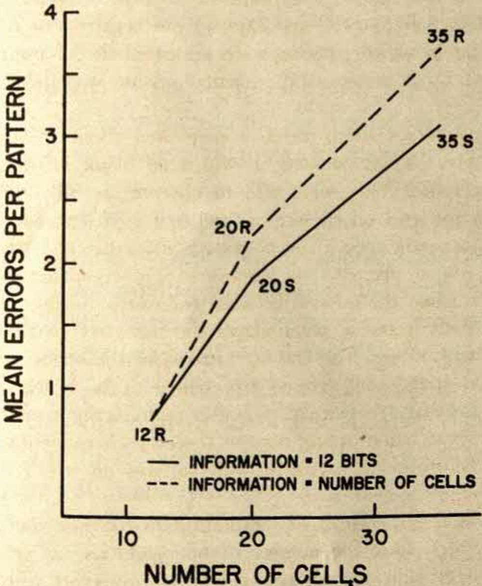


FIG. 2. IMMEDIATE REPRODUCTION: MEAN ERROR-SCORES OF THE FIVE GROUPS

when amount of information is held constant and symmetry is consequently associated with an increased number of cells: thus about 4 times as many errors are made on 35-cell symmetrical patterns as on 12-cell random patterns containing precisely the same information. The superiority of Group 20 S over 20 R and of 35 S over 35 R indicates that symmetrical patterns are easier to apprehend than random ones which occupy the same number of cells and accordingly contain more information, but these differences are not as great as one might expect.

It appears, then, that the *Ss* made only limited use of the potentialities

cedure yields values of *F* and *t* which may be over-conservative, but it contributes to our confidence that obtained differences are not due to peculiarities of the sample of patterns used. This point is of more consequence in the case of the two later experiments than at present.

for economical encoding inherent in symmetrical figures. To put the matter in very simple, concrete terms: if an *S* in Group 35 S could have restricted his attention completely to the 12 independent cells in the upper left corner of the pattern, he should have been able to reproduce these 12 cells as accurately as an *S* in Group 12 R, and to reconstruct the remaining 23 cells with equal accuracy. Some process functionally equivalent to this must have operated to produce significant differences between 20 R and 20 S, and between 35 R and 35 S, but it appears to have operated to a very limited degree. *S*s presented with symmetrical patterns clearly 'knew' that they were symmetrical; however, the responses of such *S*s conformed almost without exception to the principles of symmetry characterizing the class, whatever specific errors were made. In other words, errors were themselves almost invariably symmetrical (except, of course, when they occurred on an axis of symmetry).

An additional analysis which was somewhat incidental to the main purpose of the experiment should nevertheless be described briefly. Within Group 20 R only, the relative difficulty of patterns was determined by summating the errors made on every pattern (over all 20 cells, not merely the 12 cells previously scored) by all *S*s. This total was correlated with the number of dots (as opposed to unfilled cells) in the pattern. Separate correlation coefficients were calculated for the 'positive' and 'negative' subgroups: these were 0.67 and 0.50, respectively. This finding suggests that *S*s tended to remember *where* dots were located in the matrix, rather than *whether* successive cells were dotted. In a communication system, the amount of information which would be required to describe patterns like these in terms of *where* would increase with number of dots, and would be, on the average, considerably greater than the amount required for description in terms of *whether,* which would always be 20 bits.

Experiment II: Delayed Reproduction

One may reasonably suggest that *S* took poor advantage of the redundancy of patterns in Experiment I because of temporal limitations. Since patterns were exposed for only 1 sec., and reproduced immediately, organizational or encoding processes may not have had time to function in an efficient manner. In any case, it is well known from other situations that immediate memory has certain characteristics peculiar to itself. Accordingly, Experiment II was designed to investigate the same stimulus-variables as Experiment I under conditions involving repeated and fairly long exposure of patterns, as well as delayed reproduction. These conditions were achieved

by the employment of a paired-associates paradigm in which letters of the alphabet served as stimuli and patterns as responses.

Subjects. A total of 160 new *S*s (5 groups of 32) was obtained from the same source as before.

Materials and apparatus. The materials and apparatus were the same as used in Experiment I.

Procedure. Although the five groups were given patterns of different classes, as before, the increased difficulty of the task made it necessary to reduce the number of patterns learned by any one *S* to eight. Every group was divided into four subgroups of eight *S*s. Subgroup A was assigned the first eight patterns of the series used in Experiment I, and Subgroup C the second eight; the patterns given B and D were the 'negatives' of these given A and C, respectively. Thus a total of 16 different patterns were used (or 32, if 'positives' and 'negatives' are considered different): *i.e.* about a third as many as were scored in Experiment I.

The first eight letters of the alphabet were assigned to the eight patterns of a subgroup. Subjects were given accurate instructions as to the nature of their task. They were told that patterns designated by certain letters would be shown them, and that they would subsequently be required to reproduce every pattern as accurately as possible in response to the letter which had been paired with it. The letter designating a pattern was spoken by *E* as the pattern was presented. The duration of exposure was 10 sec., with 15 sec. intervening between the end of exposure and the presentation of the next pattern. After the 8 patterns had been exposed, *S* was given a mimeographed booklet similar to that used in Experiment I, and was given briefly the letter corresponding to the first pattern which he was to attempt to reproduce. Thirty seconds were allowed for the reproduction of the first pattern, after which *E* read the letter designating the second pattern to be reproduced, and so on until reproductions had been attempted for all eight patterns. The presentation of the patterns was then repeated, and *S* again attempted to reproduce them in response to the letters. The presentation of all eight patterns was thus alternated with the reproduction of all eight of a total of four trials. Eight random permutations of the stimuli were employed to determine separately their order in the four presentations and four reproductions. The pairing of letters with patterns remained constant, of course, throughout the four presentations. The same eight random permutations and the same assignments of letters to patterns were used for all groups: in other words, every pattern was designated by the same letter, and was exposed and reproduced in the same ordinal positions, as the 4 patterns in other classes with which it shared 12 cells.

For scoring purposes, all but the 12 common cells were ignored, as in Experiment I. The scoring of these results presented a new problem, however. In a sense, *S* was required to do two things: (a) to associate patterns with letters correctly; (b) to remember the appearance of the patterns in sufficient detail to reproduce them. It was considered entirely possible that these two tasks would not be affected in the same way by the experimental variables; therefore an attempt was made to score them separately. The frequency with which the correct response-pattern was approximated (a reproduction was considered to be 'approximately' correct if at least eight of the 12 cells were properly filled) was taken as a measure of accuracy

of associations. Accuracy of reproduction, on the other hand, was measured by scoring every reproduction against whatever pattern, out of the eight presented, it resembled most closely, with the additional restriction that no two reproductions from the same trial could be scored against the same pattern. Although these two measures were statistically independent to a considerable degree, the functions which they yielded relating the five groups were virtually identical except for differences in scale. A third measure, obtained simply by counting the number of reproductions which were imperfect in any way, *i.e.* either because one or more of the 12 critical cells were incorrectly filled or because the reproduction was not appropriate to the stimulus-letter yielded results which differed in no important respect from those of the other two. This last measure is the one which will be referred to hereafter, since it is the most obvious and straightforward of the three.

Results. An F-ratio of 41.76 ($df = 4/155$) was obtained for differences among all 5 groups. It should be remarked that homogeneity of variance does not obtain: perfect responses are so infrequent in Group 35 R and (to a lesser degree) in Group 20 R that the variance of these groups is considerably curtailed. Since this heterogeneity is itself attributable to differences in difficulty among the groups, it is not considered to invalidate the analysis, particularly in view of the high F-ratio obtained. In testing for differences among Groups 12 R, 20 S, and 35 S, however, it was considered appropriate to use an estimate of error based upon these groups alone instead of upon all five. An F equal to 12.32 ($df = 2/93$) was thus computed for the three groups over which information was held constant. A t of 4.24 was obtained between Groups 20 S and 20 R, and a t of 6.90 between Groups 35 S and 35 R. All these values are significant well beyond the 0.1-% level.

Results are plotted in Fig. 3, which is comparable to Fig. 2 of the previous experiment. It may be seen that the difficulty of symmetrical patterns is now about halfway between that of random patterns equivalent in information content and that of random patterns equivalent in number of cells; whereas in Experiment 1 symmetrical patterns were only a little easier than random patterns of equal size. These results are consistent with the hypothesis that an appreciable amount of time is necessary for the efficient encoding of redundant stimuli. It should be emphasized, however, that the handling of redundancy in the present experiment is still by no means ideal: *i.e.* the symmetrical patterns are still considerably more difficult than the 12-cell patterns, with which they are identical in informational content. The fact that there is little or no difference between Groups 20 S and 35 S is difficult to reconcile with the remainder of the data in general terms: it suggests that vertical symmetry may for some reason be easier to 'organize' efficiently than horizontal.

Experiment III: Identification

In the third experiment, Experiment III, the *S*s were required not to reproduce the patterns, as previously, but merely to learn to identify them by means of arbitrary 'names.' It was by no means obvious, on *a priori* grounds, that the results of this experiment should resemble those of the

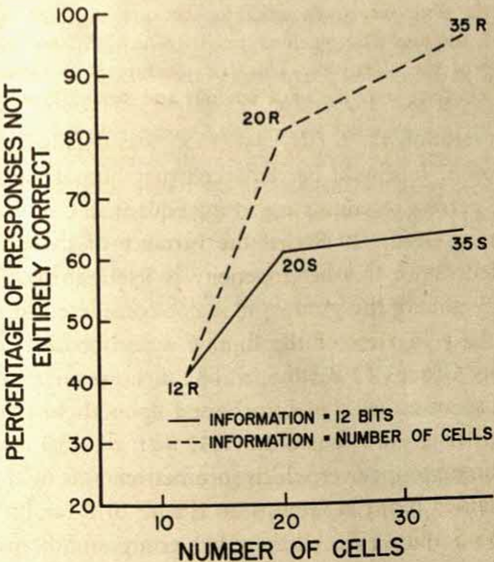


Fig. 3. Delayed Reproduction: Percentage Imperfect Responses for the Five Groups

other two. For example, it would not be unreasonable to expect that complex figures would be easier to identify than simple ones, since they contain more cues on the basis of which discrimination might be made.

Formally, the present experiment closely resembled Experiment II. A paired-associates paradigm was again used, but this time the patterns of dots were employed as stimulus-objects and letters of the alphabet as responses.

Subjects. One hundred sixty new airmen, divided into 5 groups of 30, were used.

Materials. The materials and apparatus were the same as used in Experiment I.

Procedure. With the following exceptions, the procedure was the same as in Experiment II; (a) dot-patterns were the stimulus-objects and letters of the alphabet the responses and the *S*s responded by writing the appropriate letters on an answer sheet; (b) 12 instead of 8 patterns were given each sub-group of *S*s, thus a total of 24 patterns and their 'negatives' were used for each group; and (c) an exposure-interval of 5 sec., with 15 sec. between exposures, was used on both learning- and

test-trials. The various groups and subgroups were related to one another as in Experiment II.

Scoring was accomplished simply by summating errors over all four test-trials. This involved none of the difficulties present in Experiment II since a letter given as a response was clearly either right or wrong.

An initial analysis of the results indicated that differences among the groups would be considerably less clear-cut than in the previous two experiments. Therefore a means of improving the precision of the data was sought. The use of scores on certain tests of the Airman Classification Battery (a group of tests taken by all basic airmen upon entrance into the Air Force) for purposes of co-variance control, suggested itself as an economical alternative to the collection of data from additional *S*s. Two tests, known as 'Memory for Landmarks' and 'Speed of Identification,'[6] appeared to predict a substantial portion of *S*-variance in this experiment; stanine scores on these two tests were combined by simple summation to provide a predictor variable for use in co-variance analysis.[6] In the course of this analysis it was found that a 'within-groups' correlation of -0.64 existed between the predictor variable and errors on the experimental task: thence it may be calculated that the use of co-variance control effected an increase in precision equivalent to that which would have been obtained by using more than 100 additional *S*s.

Results. By the co-variance method, an F equal to 2.72 ($df = 4/154$), significant at the 5-% level, was found for differences among the 5 groups. No differences approaching significance were found between Groups 12 R, 20 S, and 35 S; here the F was only 0.41, with $df = 2/154$. A t of 1.82 ($p = 0.1$) was obtained for the difference between groups 20 R and 20 S; and a t of 2.86 ($p < 0.01$) between 35 R and 35 S.

In Fig. 4 the mean error scores of the five groups, adjusted for differences on the predictor variable, are plotted as before. Increasing number of cells and information concomitantly again produced an increase in errors; however, this effect is considerably smaller than in either of the two reproduction experiments, whether one compares the experiments on a basis of statistical separation of groups, or in terms of proportional increase in mean error. In the present results, symmetrical patterns are much closer in difficulty to random patterns containing the same information than to random patterns containing the same number of cells. In view of the previous experiments, one may suspect that the symmetrical patterns would be significantly more difficult than those containing only 12 cells if a sufficiently large number of *S*s were employed, but whatever real difference may obtain here is certainly small. In any case, the data offer little encouragement for the view that symmetry, apart from the redundancy which it entails, has a

[6] J. P. Guilford (ed.), *Printed Classification Tests,* Washington, D.C.: U.S. Government Printing Office, 1947, Army Air Forces Aviation Psychology Program Research Report No. 5, pp. 248-251 and 376-382.

facilitating effect on memory. As in the case of reproduction, however, symmetrical patterns are significantly easier to identify than random patterns containing the same number of cells but more information.

Discussion

The results of these experiments clarify a principle of Gestalt psychology which, though at least partially correct, has heretofore been ambiguous and resistant to quantitative study. From these results it is fairly clear that one cannot make irregular patterns easier to remember by presenting them

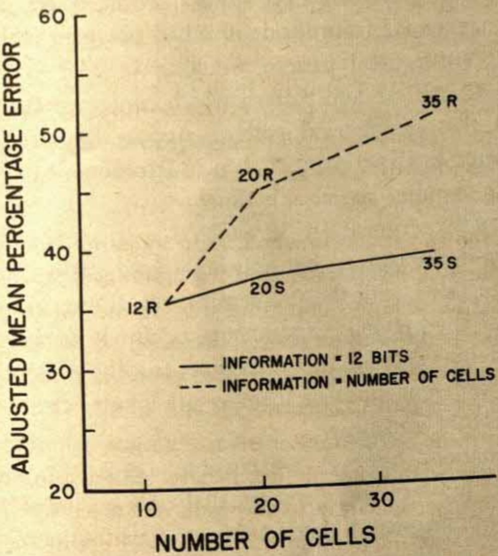


Fig. 4. Identification: Mean Percentage Error, Adjusted by Co-variance Method

in a symmetrical form—*e.g.* by placing them at right angles to a mirror to produce a symmetrical display. It should be added, in fairness to the position of the Gestalt psychologists, that they have never specifically claimed that one could. *Simplicity* and *symmetry* are both said to be factors which contribute to 'figural goodness,'[7] but in the absence of quantification it is unclear which of these factors would be expected to dominate when they are opposed to one another, as in our comparison of Groups 12 R, 20 S, and 35 S.

It is in no way surprising that random patterns are more difficult to re-

[7] Koffka, *op. cit.,* 110.

produce the more elements they contain: the fact that individuals can learn only a limited amount in a given time has never been disputed. The further finding that symmetrical patterns are more accurately reproduced than random patterns with the same number of elements (and accordingly more information) may be taken to indicate that some perceptual mechanism is capable of organizing or encoding the redundant pattern into a simpler, more compact, less redundant form. Since we know little or nothing about the nature of this mechanism, we should exercise great caution in generalizing from symmetry to other forms of redundancy which may involve different mechanisms. In the present experiments, we have seen that the process operates in a manner far from ideal: if it had operated with perfect efficiency, a 35-cell symmetrical pattern would have been reproduced as accurately as the 12-cell random pattern which contained the same information. The fact that symmetrical patterns appear to have been relatively better remembered in Experiment II than in Experiment I strongly suggests that 'organizing' requires appreciable time.

In the case of Experiment III, in which *S* was not required to store all the information in the pattern, but only enough to distinguish it from other patterns, we are faced with the problem of explaining why the results resemble those of the previous experiments as closely as they do. Such consistency would be expected if we assume (as the Gestalt psychologists seem to) that different kinds of memory all involve the storage of traces in approximately the same way.

Other explanations, however, are possible. Let us suppose that the patterns are made up of certain elements, that only a limited number of elements can be apprehended or abstracted from a pattern on a given presentation, and that the elements abstracted are either associated with the correct response, in the case of a learning trial, or else constitute the effective stimulus-object to which a response is made, in the case of a test-trial. The nature of these elements need not be specified, if a positive correlation between number of elements and number of cells is granted. Now it is clear that the probability of correct response will be a function of the probability that identical elements will be 'drawn' on learning trials and on subsequent test-trials, and that the latter will in turn depend upon the proportion of elements which can be apprehended from a given presentation. The more elements a pattern contains, the greater is the chance that the elements abstracted will differ from learning- to test-trial, and hence the smaller the chance of a correct identification. If some of the elements within a pattern are identical with one another, as in the case of a symmetrical or otherwise

redundant pattern, the probability of abstracting identical elements on learning- and test-trials, and hence the probability of correct response, will be increased. These relationships would hold, however, only for patterns containing more elements than could be apprehended at once, and would not apply to simpler patterns.[8]

French, working in this Laboratory, and using the same apparatus, has previously studied the difficulty of learning to identify patterns as a function of their complexity.[9] Varying the number of randomly chosen dots in a matrix of 112 cells from 1 to 12, for separate groups, he found that patterns containing about 7 dots were identified most accurately, with errors increasing on either side of this minimum. He further discovered that the easiest patterns at any given level of complexity tended to be characterized either by a fortuitous approximation of symmetry, or by some degree of 'good continuation,' or 'stringing-out' of points. An aspect of French's results which is at first puzzling (from the point of view of the preceding paragraph) is that symmetry and good continuation were found to be favorable to identification even in the case of patterns containing fewer than the optimal number of dots. These characteristics occurred only rarely among his patterns, however, and may have had their chief value as *cues* to the identification of the occasional patterns possessing them.

Adams *et al.* have recently reported that *S*s learn to identify symmetrical shapes more easily than asymmetrical shapes when the two are matched with respect to number of lines and angles.[10] Although symmetrical and asymmetrical shapes were mixed in the same list, they occurred in equal number; hence it is unlikely that symmetry had any more value *as a cue* than did asymmetry. These authors suggest that symmetry may be considered a form of redundancy, but do not develop the idea further. It seems fairly clear that the difference between their symmetrical and asymmetrical shapes was comparable to the difference between our Classes 20 S and 20 R, or 35 S and 35 R.

The usefulness of concepts derived from information-theory in the quan-

---

[8] The above interpretation is compatible with recent developments of learning theory in statistical terms. See especially W. K. Estes, Toward a statistical theory of learning, *Psychol. Rev.*, 57, 1950, 94-107; W. K. Estes and C. J. Burke, A theory of stimulus variability in learning, *ibid.*, 60, 1953, 276-286; R. R. Bush and F. Mosteller, A mathematical model for simple learning, *ibid.*, 58, 1951, 313-323; A model for stimulus generalization and discrimination, *idem.*, 413-423.

[9] French, Identification of dot patterns from memory as a function of complexity, *J. Exper. Psychol.*, 47, 1954, 22-26.

[10] O. S. Adams, P. M. Fitts, M. Rappaport, and M. Weinstein, Relations among some measures of pattern discriminability, *J. Exper. Psychol.*, 48, 1954, 81-88.

tative investigation of a concept as vague as that of the 'good figure' has been demonstrated by the experiments reported here. It should be emphasized, however, that information-theory is not psychological theory, and hence not competitive with any particular psychological theory. Whether the results of these studies may best be accounted for in terms of associative bonds, or cerebral fields, or some other set of mechanisms, remains an open question.

SUMMARY

Three experiments investigating the effect of symmetry on memory for patterns are reported. In Experiment I, *S* was required to reproduce patterns immediately after brief exposure. In Experiment II, patterns were used as response-members of paired associates, and duration of exposure, interval between exposure and reproduction, and number of exposures were all greater than in Experiment I. In Experiment III, *S* was not required to reproduce the patterns at all, but instead to identify them: *i.e.* the patterns were used as stimulus-members of paired associates.

The stimulus-objects were patterns of dots presented in rectangular matrices. In all three experiments they were varied in such a way as to permit comparison between memory for symmetrical patterns and memory for (a) asymmetrical patterns with the same informational content—hence fewer cells, and (b) asymmetrical patterns occupying the same number of cells, hence containing more information.

In none of these experiments were symmetrical patterns easier to remember than asymmetrical patterns with the same informational content; in Experiments I and II they were significantly more difficult. In all of them symmetrical patterns were remembered more easily than asymmetrical patterns occupying the same number of cells; though this superiority was least striking in the case of immediate reproduction. Random patterns were found in all cases to be more difficult the greater their complexity, *i.e.* the more cells they occupied.

These results are believed to constitute an important clarification of the Gestalt doctrine that 'figural goodness' is favorable to memory.

THE EFFECT OF TILT ON THE VISUAL PERCEPTION OF PARALLELNESS

By Alan M. Rochlin, Duke University

This experiment is part of a program of research designed to analyze the visual perception of spatial relationships. Earlier experiments in the series have yielded functional dependencies for several simple spatial structures. In a study by Salomon, on the ability of *S*s to align a dot in such a manner that it appears to fall on an extrapolated extension of a tilted reference line, a linear relation was found between accuracy of setting and line-to-dot distance, the slope of the function varying inversely with the length of line.[1] Sulzer demonstrated the importance of essentially the same stimulus-determinants—line-length and inter-line distance—for the perception of parallelism.[2] In both studies, evidence of the importance of yet a third stimulus-determinant—tilt of line—was found. The rôle of tilt in the perception of parallelness was investigated in the present experiment.

Method

The apparatus, procedure, and general conditions of the present experiment were with but few exceptions the same as those of Sulzer.[3] Those modifications which have been made will be explicitly mentioned.

Visual situation and task. The visual field consisted of a brightly lighted (15.2 m. lam.) area of translucent glass bounded by the circular aperture of a blackened tube. This tube projected from *O*'s eyes to a point 6 in. away from the lighted glass. The lighted field was 17 in. in diameter and was viewed from a distance of 30 in., forming a visual angle of approximately 32.° Two thin black lines of equal length and thickness, in the form of a rectangle with two opposite sides missing, appeared in the center of the field approximately parallel to each other. It was the task of *O*, by turning a uniform, circular disk, to tilt the adjustable line about its center until the two lines appeared 'exactly parallel,' as required by the instructions.

* Accepted for publication June 14, 1954. This paper was adapted from a doctoral dissertation submitted to Duke University. The writer is indebted to Professor Karl Zener for assistance throughout the course of the research.

[1] A. D. Salomon, Visual field factors in the perception of direction, this Journal, 60, 1947, 68-88. Salomon's findings were confirmed in a later study by J. D. Miller and B. Waldron (The extrapolation of a straight line, Honors Thesis, Mount Holyoke College Library, 1951.)

[2] R. L. Sulzer, A determination of several functions relating sensitivity of perception of parallelness to stimulus dimensions, Doctoral Dissertation, Duke University Library, 1954, 1-57. See R. L. Sulzer and K. E. Zener, A quantitative analysis of relations between stimulus determinants and sensitivity of the visual perception of parallelness, *Amer. Psychol.*, 8, 1953, 444.

[3] For a detailed account, including a more elaborate description of the calibration of the apparatus and preparation of stimulus-materials, see Sulzer, *op. cit.*, 4-10.

Modification of the apparatus. The major modification of the apparatus employed by Sulzer was a change in the sensitivity of the variable line to movement of the adjusting disk. This change was accomplished mechanically by a reduction in gear-ratio in the drive mechanism from 100:1 to 40:1. While before modification the least perceptible movement of the adjusting disk, 1° of arc, caused the tilt of the adjustable line to be changed 9″, after modification the same movement caused a rotation of 22.5″. This modification served to minimize the effects of motor sets, to reduce the time required for each response, and to eliminate extraneous cues produced by the mechanical operation of the apparatus, such as auditory cues arising from the action of the gears. A second major change involved a change in the material used as a screen to shield the glass area from *O* between trials. Instead of a black cloth, translucent, grained lucite was used, which permitted the light-intensity to be kept relatively constant whether the screen was up or down.

Stimulus-determinants: (1) Tilt. Twelve angular positions of the parallel lines were used, ranging from 0° (horizontal) to 165° in 15° steps. These tilts were measured in a clockwise direction from the horizontal, *i.e.* the 15° line slanted from *O*'s lower right to upper left. For convenience, the tilts of 15°, 30°, 45°, 60°, and 75° were designated as 'left-ascending' and the tilts of 105°, 120°, 135°, 150°, and 165° as 'right-ascending.' The adjustable line was placed in the center of the visual field and was rotated about its midpoint. It was always placed above the fixed line in all tilts except the vertical where it was to the right of the fixed line.

(2) Line-length. Only two line-lengths were used in this study to maximize the number of angular positions tested. Two representatives line-lengths were selected on the basis of an analysis of Sulzer's data on the horizontal and vertical positions, and on the basis of a pilot study carried out with the 135° position. Three *O*'s were used in the pilot study; the apparatus and procedure were the same as Sulzer's. As in Sulzer's data for the horizontal and vertical positions, the line-lengths seemed to fall into three apparent groups. The ⅛-in. line had the greatest variance; the ½-, 1-, and 2-in. lines had the least; and the ¼-in. line fell clearly between those groups with respect to variance. It was decided to select, therefore, the ¼-in. line (to avoid the extreme variability produced by the ⅛-in. line) and the 1-in. line (the middle one of the group with least variability).

(3) Inter-line distance. The four inter-line distances selected were ¼, ½, 1, and 2. The ¼-in. inter-line distance was used as the smallest distance separating the two lines, since in Sulzer's study line-separations of 1/32 and ⅛ in. yielded relatively flat curves when variance was plotted against tilt, showing for these small distances a minimum, if any, effect of tilt upon sensitivity to parallelness.

Observers. Four college women (*M*, *S*, *W*, and *Y*) served as paid *O*s. They were psychologically naïve, and care was taken to avoid revealing either the purpose of the experiment or the fact that variability of setting was of interest. The *O*'s were selected on the bases of the following criteria: (1) Visual tests; all *O*s had visual acuities of 20/20 or better (3 of the 4 with correction) as determined by the Snellen Acuity Chart and the Bausch and Lomb Ortho-rater; they were normal on depth, phoria, and color-tests, and gave no indication of an astigmatic condition in tests with the Lancaster-Regan Astigmatic Dial.[4] (2) Tests of dominance: all the *O*s

[4] In addition to these tests, all *O*s were required to undergo a complete visual examination. The writer is indebted to Dr. W. B. Anderson, of the Department of Ophthalmology of Duke Hospital, whose thorough examination revealed no pathology of vision for any of the *O*s.

were strongly right-handed as determined by interview and by a standard laboratory handedness-questionary. The *O*s were also selected on the basis of a test for left cerebral dominance in vision, adapted from Jasper and Raney.[5]

Training. Since Sulzer's investigation of tilt was a supplement to his main experiment which dealt only with lines in the main axes of space, his *O*s experienced the horizontal and vertical positions for a prolonged period of time, and then were tested without preliminary training for the effect of angular position of the lines. To determine the extent to which the differences in variability for different tilts obtained by Sulzer were a function of differential training, a pilot study was performed. The three *O*s mentioned previously were introduced to the task for the first time in this pilot study. The procedure and apparatus were identical with Sulzer's except for the number of settings per condition—only 20 settings rather than 40 were used. Each day, two lengths of line, ¼ and 1 in., were tested at two inter-line distances, ¼ and 2 in. The *O*s experienced all of these conditions for lines in the vertical position on Days 1 and 2. On Days 3 through 7, all four of these conditions were presented at a tilt of 135°, the position of maximal variability in Sulzer's study. On Days 8 and 9, the vertical position was reintroduced to note possible interaction-effects.

TABLE I

AVERAGE VARIANCE OF SETTINGS (IN DEGREES OF LINE-TILT) FOR TWO LINE-LENGTHS (*L*) AT TWO INTER-LINE DISTANCES (*D*) FOR VERTICAL LINES

L	*D*	Day			
		1	2	8	9
¼ in.	¼ in.	.0742	.0361	.0854	.0612
¼ in.	2 in.	.0658	.0480	.1779	.0617
1 in.	¼ in.	.0297	.0197	.0221	.0273
1 in.	2 in.	.0494	.0507	.0561	.0488

In Table I, variance of settings is presented for the two line-lengths at the two inter-line distances at the vertical position for Days 1, 2, 8, and 9. At Day 2, there was a drop in variability. When the vertical was reintroduced at Day 8, however, there was a marked increase in variance over Day 2, which decreased on Day 9, indicating an interference-effect between the two angular positions. The five-day period for the 135° lines showed no consistent pattern.

In the main experiment, therefore, to allow the *O*s a maximal amount of adaptation to the task and to minimize possible differential effect of training, the following training procedure was carried out: The *O*s were first presented with ½-in. line-lengths (a length intermediate to the ¼ and 1-in. lengths), at the four major angular positions, 0°, 90°, 45°, and 135°, and at two inter-line distances, ¼ and 2 in. They made 20 settings for each condition which was encountered twice. This first training-sequence lasted four days. During the next four days, immediately preceding the start of experimental Day 1, they experienced the same conditions, but they were required to make 40 settings per condition.[6]

[5] H. H. Jasper and E. T. Raney, The phi test of lateral dominance, this JOURNAL, 49, 1937, 450-457.

[6] One of the four *O*s, because of certain delays arising from the selection procedure, experienced only the second training sequence.

Experimental design. Each *O* was tested once under each of the 96 conditions (12 tilts 2 line-lengths × 4 inter-line distances). The work required a total of 26 experimental days. For each of the first eight days, the *O*s experienced three conditions, a particular tilt for one length of line at three inter-line distances. On days following the first, work would be completed for the remaining inter-line distance and a new tilt would be presented. From experimental Day 9 throughout the remainder of the experiment, four conditions were presented each day, *i.e.* all four of the inter-line distances for a given line-length at a given tilt.

The order of presentation of the 96 conditions was designed to minimize the development of sets to any of the determinants and the differential effects of learning and motivation. To minimize the effects of sets for angular position, those lines on or near the main axes were alternated with lines in oblique positions for each *O*.[7] Further, the number of consecutive right or left obliques was limited to two. This condition was balanced among *O*s by interchanging for alternate *O*s the left and right ascending tilts of the lines. On any given day, each *O* experienced a line-tilt different from those which the others experienced, a procedure intended to reduce position-effects over the sequence of conditions.

With regard to inter-line distance, two *O*s followed a regular pattern in which the smallest inter-line distance did not appear just before or after the largest. The remaining two *O*s had a randomized presentation of inter-line distances. To minimize sets for line-length, no *O* experienced more than two consecutive lines of the same length for differing tilts, *i.e.* no *O* experienced the same line-length for more than eight consecutive conditions. For a given tilt, two *O*s experienced one of the line-lengths first, while the other two saw the other line-length first.

To minimize differential effects of learning and motivation, no tilt was presented for a second line-length before all the other tilts had been experienced. After the first 12 tilts had been presented for a particular line-length, the second 12, for the remaining line-lengths, were presented in an order which was the mirror-image of the first.[8]

Treatment of data. Two statistical measures were computed for the 40 settings obtained under each condition. They were the mean (M) and the variance (SD^2). M is here to be considered a measure of constant error, for it specifies the difference in tilt between the objectively correct setting and the average setting of each *O*. SD^2 is here to be considered an inverse measure of perceptual sensitivity, for it specifies the sum of the squared deviations of each *O*'s setting from his own computed mean-setting. Variance was selected as a measure of dispersion rather than the more commonly employed SD because of the additive property of variance, which permitted analysis of the results by certain pooling techniques.

The adjustable line was variably offset from the objective parallel position (scale value = 0) by *E* prior each setting. The direction of offset was alternated, first by a clockwise movement of the adjustable line, and then by a counter-clockwise movement. M and SD^2 were computed separately for each direction of offset, and from

[7] The angular positions 0°, 15°, 75°, 90°, 105°, and 165° were considered to be on or near the main axes, and were alternated with 30°, 45°, 60°, 120°, 135°, and 150°, which were considered oblique.

[8] After the experiment was completed, the *O*s were questioned regarding their general attitude toward the task. They reported that they had enjoyed doing it and that they had experienced no decrease in motivation.

these the average M and SD^2 for the group of settings for a particular condition were obtained.

RESULTS

The results of the experimental variations of the three major stimulus-dimensions—tilt, line-length, and inter-line distance—will be presented in three major sections: the first will deal with the general effects of tilt upon variance; the second with the effects of tilt upon the relation of inter-line distance to variance; and the third with the effects of the experimental conditions upon constant error.

(1) General effects of tilt upon variance. In Figs. 1 and 2 the average variances for each inter-line distance are plotted separately by the ¼- and 1-in. lines, respectively. These plots show that the least amount of variability

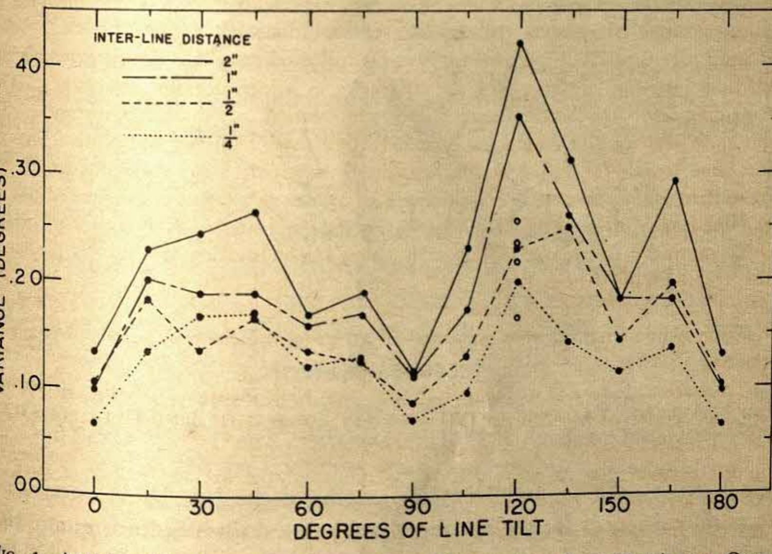


FIG. 1. AVERAGE VARIANCE OF SETTINGS (IN DEGREES OF LINE-TILT) FOR ONE-QUARTER INCH LINES AT FOUR INTER-LINE DISTANCES AND TWELVE LINE-TILTS

consistently occurs at the horizontal and vertical positions for both line-lengths and for all inter-line distances.

For all tilts and line-lengths, variance increases with inter-line distance. Confirming the consistency of this function is the fact that, out of the 72 possible comparisons between adjacent inter-line distances for all tilts and both line lengths, there are a total of only 8 inversions, all minor. Of these, there are two inversions for the right-ascending tilts; four for the left-ascending tilts; and two for the lines on the main axes.

Comparison of Figs. 1 and 2 shows that the ¼-in. line-length produced greater variability than the 1-in. line-length. Of the 48 possible comparisons between the two line-lengths, there are only 2 inversions, both occurring at the 2-in. inter-line distance for line tilts 60° and 150°.[9]

These average data show greater variability than those obtained by Sulzer for comparable line-lengths and inter-line distances in a study dealing only with lines in the horizontal and vertical positions. In all of the 16 possible

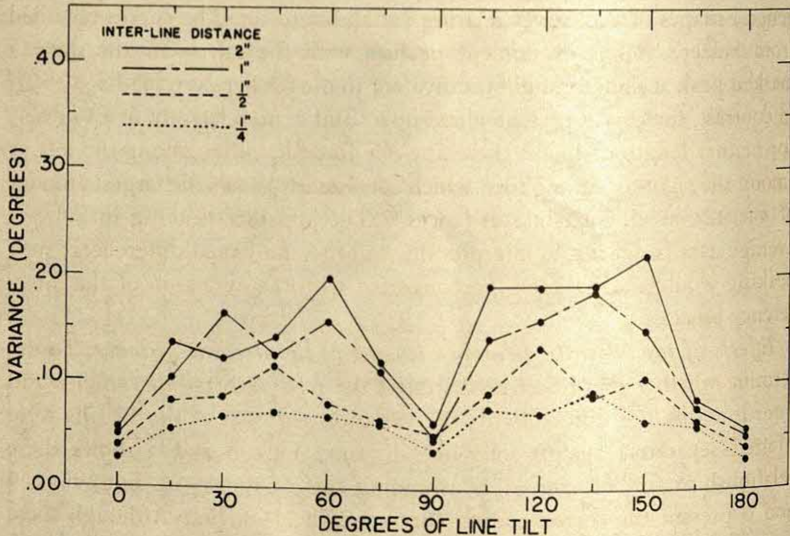


FIG. 2. AVERAGE VARIANCE OF SETTINGS (IN DEGREES OF LINE-TILT) FOR ONE-INCH LINES AT FOUR INTER-LINE DISTANCES AND TWELVE LINE-TILTS

comparisons (the ¼ and 1-in. line-lengths at the ¼-, ½-, 1-, and 2-in. inter-line distances for the horizontal and vertical tilts) Sulzer's group data have smaller variances than the group in the present study. It seems unlikely that these differences are due to differences in apparatus. This possibility was checked in the following two ways: *O*s from the pilot studies, mentioned previously, who had experienced the apparatus before modification, were tested after, for the same stimulus-conditions: their variances

[9] Although the individual results in the main show the general trends, certain individual differences do emerge. Two *O*s show no general consistency with regard to the effect of increasing inter-line distance upon variance. This is true of *S* for both lengths of line; of *W*, for only the ¼-in. line-length. *M*, for only the ¼-in. line-length, shows much less of the differential effects in variance between lines on the main axes and all other line-tilts. Finally, one *O*(*Y*) shows no consistent difference in variability between line-lengths for tilts ascending to the right.

were slightly lower, when tested after modification. Two *O*s who had never experienced the task were tested for lines in the vertical position after the apparatus was modified; their variances were comparable to those obtained by Sulzer. The higher variances obtained in this present study probably result from interference produced by the order of presentation of the different tilts. The major evidence for such an interference-effect has been mentioned in the previous section (Table I).

A second difference between the data of the two studies concerns the general shapes of the curves relating variability to tilt. The curves obtained from Sulzer's minor experiment dealing with the effects of tilt show a marked peak at the 45° angle (equivalent to the tilt here specified as 135°). In contrast, the curves presented in Figs. 1 and 2 are relatively flat. Further, for neither length of line is there any one line-tilt, either among the left or among the right-ascending tilts, which consistently shows the largest amount of variance at all inter-line distances.[10] The apparent peaking in Sulzer's average data is difficult to interpret due to gross individual differences; such peaking would, nevertheless, be consistent with the operation of the interference-process.

Effects of tilt upon the relation of variance to inter-line distance. To determine whether the different tilts effect the relation between variance and inter-line distance differentially, the variances for each of the 12 tilts were plotted separately against inter-line distance. Figs. 3 and 4 shows these relationships for left and right-ascending tilts of the ¼-in. line; Figs. 5 and 6 present the corresponding curves for the 1-in. line. Although these curves reaffirm the absolute differences in variance between lines on the main axes and all other line-tilts, there appears to be no consistent relation among the curves for the line tilted away from the main axes. Comparison of the two line-lengths indicates that the 1-in. length produced a more compact group of curves for line-tilts off the main axes.

Since many curves are presented in these figures, possible differences in slope between the lines on the main axes and all other line-tilts are not

---

[10] The sharp peak at 120° for the ¼-in. line deserves special comment. It is felt that these points are too high, since the extreme variance was brought about largely by one *O* (*T*) who experienced the ¼-in line for the first time at this tilt and reported after the experiment was completed that she had a great deal of trouble dealing with this particular tilt. Apparent differences in variance between left and right-ascending tilts for the ¼-in. line may be considered due in part to this extreme variability by the one *O* at this tilt. The average variances for the three *O*s (excluding *T*) are presented in Fig. 1 for the test-conditions at the 120° line-tilt. In the order of increasing variance, these points represent the ¼, ½, 1, and 2-in. inter-line distances.

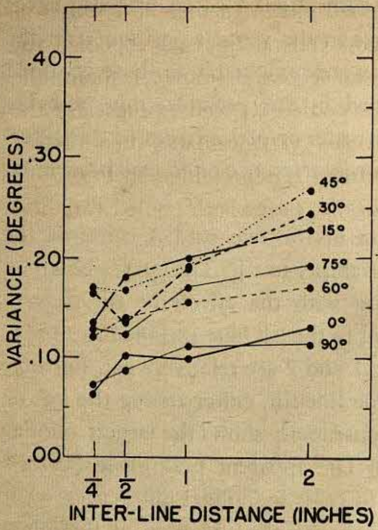


FIG. 3. AVERAGE VARIANCE OF SETTINGS (IN DEGREES OF LINE-TILT) PLOTTED FOR EACH OF THE LEFT-ASCENDING TILTS AND THE HORIZONTAL AND VERTICAL TILTS FOR ONE-QUARTER INCH LINE-LENGTHS

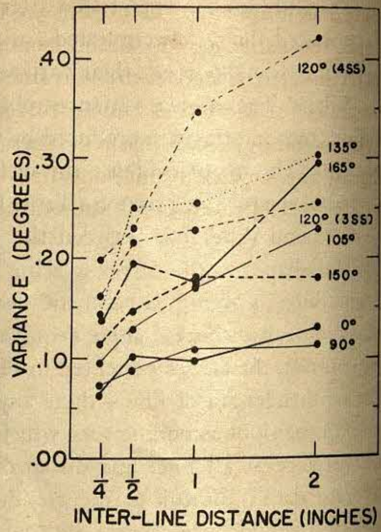


FIG. 4. AVERAGE VARIANCE OF SETTINGS (IN DEGREES OF LINE-TILT) PLOTTED FOR EACH OF THE RIGHT-ASCENDING TILTS AND THE HORIZONTAL AND VERTICAL TILTS FOR ONE-QUARTER INCH LINE-LENGTHS

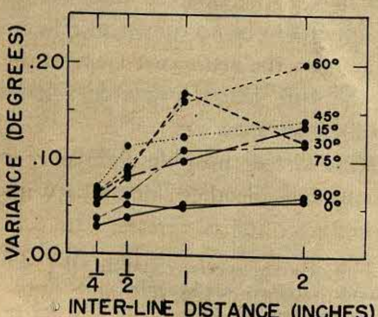


FIG. 5. AVERAGE VARIANCE OF SETTINGS (IN DEGREES OF LINE-TILT) PLOTTED FOR EACH OF THE LEFT-ASCENDING TILTS AND THE HORIZONTAL AND VERTICAL TILTS FOR ONE-INCH LINE-LENGTHS

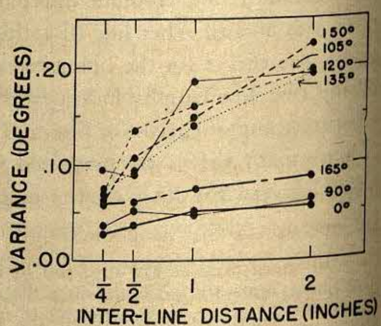


FIG. 6. AVERAGE VARIANCE OF SETTINGS (IN DEGREES OF LINE-TILT) PLOTTED FOR EACH OF THE RIGHT-ASCENDING TILTS AND THE HORIZONTAL AND VERTICAL TILTS FOR ONE-INCH LINE-LENGTHS

immediately apparent. To present a simplified picture of these differences in slope, the variances for the horizontal and vertical lines were pooled for each line-length and each inter-line distance. In a similar fashion, the variances for the left-ascending tilts were pooled, as well as those for the right-ascending tilts.[11] Each of these pooled variances was then plotted against the square-root of inter-line distance, as represented in Fig. 7, to rectify the functions. Fig. 7 shows that the slopes for both pooled variances

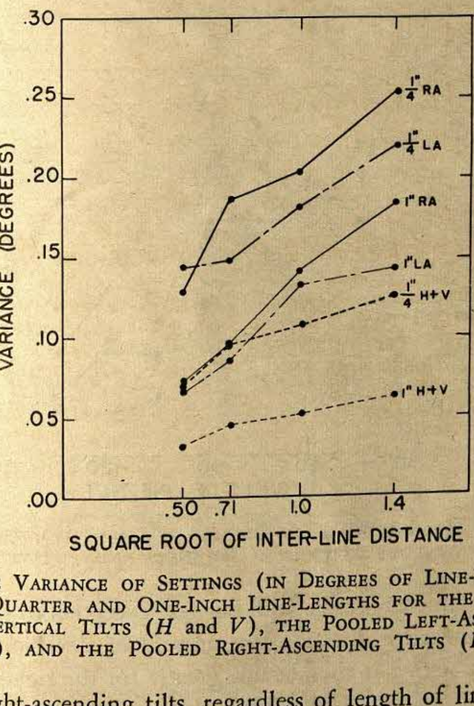


FIG. 7. AVERAGE VARIANCE OF SETTINGS (IN DEGREES OF LINE-TILT) PLOTTED FOR THE ONE-QUARTER AND ONE-INCH LINE-LENGTHS FOR THE POOLED HORIZONTAL AND VERTICAL TILTS (*H* and *V*), THE POOLED LEFT-ASCENDING TILTS (*LA*), AND THE POOLED RIGHT-ASCENDING TILTS (*RA*)

of left and right-ascending tilts, regardless of length of line, are steeper than those for the pooled variances of the lines on the main axes. With but one exception, the pooled right-ascending tilts have greater variances than the pooled left-ascending tilts. There is, however, no indication of differences in slope among the curves representing the left and right-ascending tilts, nor between the curves for the different line-lengths.

[11] In computing the pooled variances for the right-ascending tilts for the ¼-in. line-length, the data of only three of the *O*s were used for the 120° line-tilt. The reason for the exclusion of the one *O* was to prevent undue biasing of the curve representing the pooled right-ascending tilts by the extremely high variances of that *O*.

Effect of tilt upon constant error. Figs. 8 and 9 represent the mean constant errors plotted against line-tilt for the four inter-line distances (¼- and 1-in. line-lengths, respectively). For all tilts except the vertical, a negative sign indicates an average setting with the right ends of the lines too close together.[12] For vertical lines a negative sign indicates that the lower ends of the lines are too close together. Two relations immediately become apparent: the direction of the offsets from the objectively correct

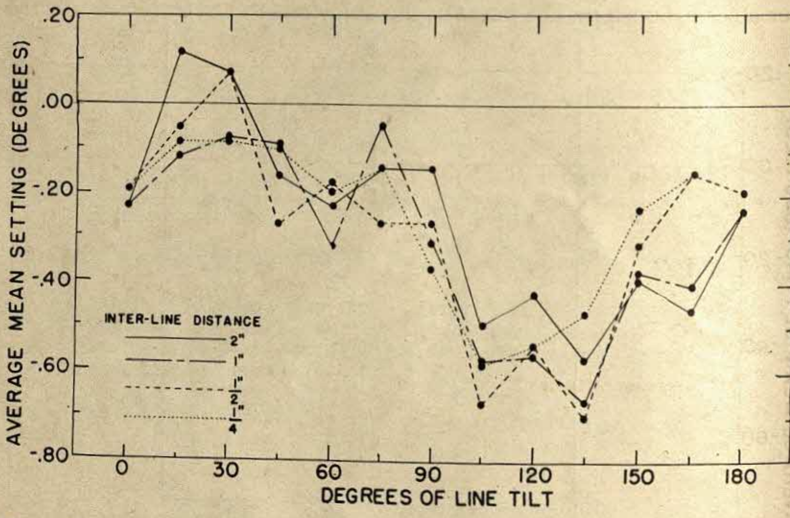


FIG. 8. AVERAGE MEAN SETTING (IN DEGREES OF LINE-TILT) FOR ONE-QUARTER INCH LINES AT FOUR INTER-LINE DISTANCES AND TWELVE LINE-TILTS

setting for both line-lengths is, with but few exceptions, in what has been defined as the negative direction; the amount of deviation from the objectively correct setting is considerably greater for the right-ascending tilts than for the left-ascending tilts. There appears to be no general relation distinguishing one inter-line distance from another for the constant-error measures.

The greatest increase in constant error occurs consistently between 90° and 105°, that is, with lines whose upper ends tilt slightly away from the vertical to the right. This increase differs markedly from those between lines on the main axes and the other tilts adjacent to them, and is to be contrasted particularly with the change between 75° and 90°, that is, with lines

[12] A clockwise movement of the adjustable top line was termed a negative movement; a counter-clockwise movement was termed a positive movement.

whose upper ends tilt slightly away from the vertical to the left. These findings are remarkably consistent with those for individual *O*s.[13]

Discussion

This experiment was conducted to investigate two aspects of the effect of tilt on the visual perception of parallelism. The first concerns the general quantitative effects produced by tilt; the second concerns the effect of tilt upon the functions already established in the perception of parallelism for other parameters. The intent of this study was to vary systematically the

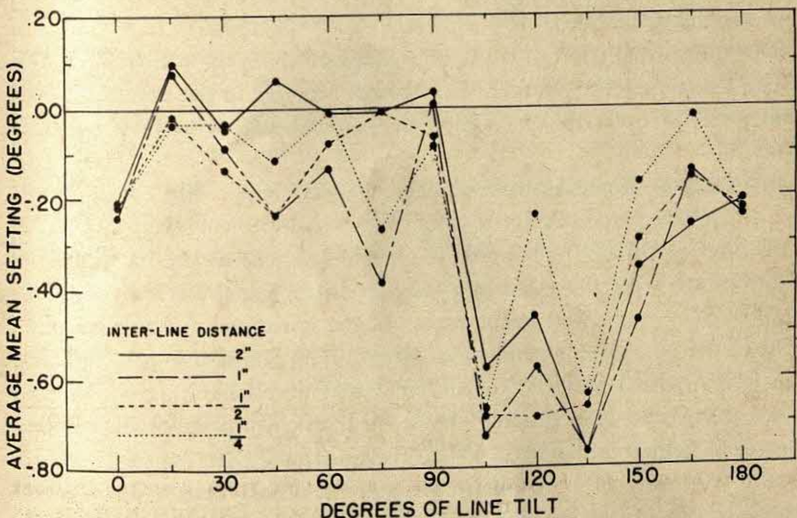


Fig. 9. Average Mean Setting (in Degrees of Line-Tilt) for One-Inch Lines at Four Inter-Line Distances and Twelve Line-Tilts

tilt of the visual stimulus-object under different conditions of line-length and inter-line distance, with the rôle of the visual framework minimized, and the position of *O* held constant. Under these conditions, the effects of the line-tilts upon variability and constant error of adjustment are attributable to two primary kinds of factor governing the perception of the main axes of space: (1) visual factors—the position of the lines on the retina and the resulting locus of cortical excitation in the visual areas; (2) nonvisual factors—such as vestibular and kinaesthetic cues.

[13] The one exception occurs for *S*, with the ¼-in. line-length. For this *O*, at this line-length, the amount of deviation from the objectivity correct setting is greater for the left-ascending tilts, with the greatest increase in constant error occurring between 75° and 90°.

General effects of tilt. The data of this experiment show that for all conditions the least amount of variability consistently occurs for lines in the horizontal and vertical positions. It is known from Sulzer's study of tilt, however, that such differential variability is dependent upon the amount of separation between the two lines to be adjusted as parallel. Sulzer found, for the 1/32- and ⅛-in. inter-line distances and with a 1-in. line-length, almost no differential effect of tilt upon variance. Since in the present experiment tilt yielded definite effects with a ¼-in. inter-line distance, the inter-line distance at which tilt begins to be effective is between ⅛- and ¼-in. for the 1-in. line-length and at least below the ¼-in. inter-line distance for the ¼-in. line.

This differential effect of tilt upon perception is not unique to the setting of one line parallel to another. Two studies reported in the literature show similar effects upon the perception of perpendicularity and upon visual acuity.

In a study at Mount Holyoke College, *O*s were required to adjust a variable line to be perpendicular to a baseline.[14] Adjustments were studied as a function of tilt of the baseline. Eighty inclinations of the baseline over a 360° range were presented in a random order. Variability (measured in terms of *SD*) was lowest in the region of the horizontal and vertical positions of the objective perpendicular. Higgins and Stultz, using the parallel-line test of visual acuity at ten different magnifications, have shown that visual acuity is between 10–20%, lower for lines passing diagonally through the visual field at an angle of 45° to the horizontal than for parallel lines passing vertically or horizontally through the visual field.[15] They found that perceptibility of the 45° lines was improved, while perceptibility of lines on the main axes of space diminished, when *O* tilted his head. These studies indicate that a differential effect is operative in the perception of several simple spatial relational structures as well as for visual acuity when the stimulus-items are tilted off the main axes of space. It would be premature, at present, to attempt to relate these findings to structural differences in the visual system, since there is a paucity of information in this area, particularly for situations allowing free fixation.

The obtained differences between perceptual sensitivity for stimulus-items in oblique positions and the horizontal and vertical positions may in part

---

[14] Psychophysical Research Summary Report (1946-1952), Psychophysical Research Unit Mount Holyoke College, NAVEXOS P-1104, Special Devices Center, Technical Report, No. SDC-131-1-5, 1953, 151-155.

[15] G. C. Higgins and K. Stultz, Visual acuity as measured with various orientations of a parallel-line test object, *J. Opt. Soc. Amer.*, 38, 1948, 756-758.

be due to differential effects of past experience outside the experimental situation. In training the *O*s for the present experiment, care was taken to insure that the *O* experience lines tilted off the main axes as often as they experienced lines on the main axes. It was realized at the time, however, that the greatest amount of perceptual experience with spatial relations occurs in life-situations with structures in the horizontal and vertical positions. Behavior is, in general, oriented in relation to the main axes of space, and sensitivity of perception may, therefore, be expected to be greater for objects located in line with these axes than for objects in oblique positions.

The data of this experiment have shown that for both measures of perceptual activity there is a difference between the left and right-ascending tilts. Right-ascending tilts produced greater constant errors and variability than the left-ascending tilts. In absolute amount, the differences between left and right-ascending tilts were greater for the constant-error measures than for the variance measures, indicating that the consistency of setting is less effected by the direction of tilt than is the accuracy of setting.

It must be emphasized that these differences for direction of tilt were found in *O*s selected for right-hand dominance and left-cerebral dominance in vision.

To determine whether the differences obtained between direction of line-tilt are related to a factor of dominance, a second group of *O*s, selected this time for left hand-dominance and right cerebral dominance in vision, should be tested. It should be pointed out, however, that if laterality is a factor in producing the differences obtained between left- and right-ascending tilts, the prediction of opposite results would not necessarily follow. Rather it is suggested that opposite results or no difference between lines ascending to the left or right would be found. The latter possibility must be considered in view of the fact that the left handed individual lives in an environment constructed to meet the needs of the right handed.

Effect of tilt upon other parameters. A second purpose of this study was to determine whether variations in tilt of the visual stimulus-objects effect the relation of line-length and inter-line distance to the perception of parallelism. The results indicate that, with but two exceptions, the ¼-in. line-length produces greater variability than the 1-in. length, and that, with only 8 minor exceptions out of 72 possible comparisons, variance increases as a function of inter-line distance.

To maximize the number of line-tilts studied, the number of line-lengths used was limited to two. This small number provides an inadequate basis

to determine whether the relations between variance of setting and line-length, established by Sulzer for the horizontal and vertical positions, are radically changed. The absolute increase in variance for the lines tilted off the main axes of space indicate that at least the intercept-constants for Sulzer's empirically fitted curves would have to be changed.

That both the intercept and slope-constants obtained by Sulzer for empirically fitted curves relating inter-line distance and variance of settings for lines on the main axes would have to be altered is indicated by the curves presented in Fig. 7, which show that the curves representing the pooled variances of the left-ascending tilts and the pooled variances of the right-ascending tilts have not only greater variances but also steeper slopes than the pooled variances of the horizontal and vertical lines. It is of interest in light of the previous discussion of direction of tilt, that there appears to be no difference in slope for lines ascending to the left and lines ascending to the right.

These changes, produced by line-tilt upon the functions relating the sensitivity of the perception to parallelness to line-length and inter-line distance, suggest the importance of studying the effects of other simple determinants, such as level of illumination, contour of the visual field, and orientation of *O* to the main axes of space.

SUMMARY

Two thin black lines were presented in the center of an otherwise homogeneously illuminated circular field. *O* adjusted one line about its center until the two lines appeared parallel. The effects of three principal experimental variables—line-tilt, line-length, and inter-line distance—were investigated in 4 *O*s for two measures of perceptual activity—variance of settings and constant error of settings. Twelve tilts of line, two line-lengths, and four inter-line distances were tested. For all conditions, oblique lines yielded greater variance of settings than did horizontal and vertical lines. Right-ascending tilts produced greater variance of settings and decidedly greater constant errors than did left-ascending tilts. The effects of differences in line-length and in inter-line distance were consistent with earlier findings. The slope of the functions relating inter-line distance to variance of settings was steeper for both the pooled left and right-ascending tilts than for the pooled horizontal and vertical lines.

SIMULTANEOUS AND SUCCESSIVE DISCRIMINATION UNDER IDENTICAL STIMULATING CONDITIONS

By M. E. BITTERMAN, D. W. TYLER, and CLAUDE B. ELAM
University of Texas

A number of recent reports from this laboratory have been concerned with the relative difficulty of simultaneous and successive discrimination under a variety of experimental conditions.[1] The designations *simultaneous* and *successive* derive from the fact that the two types of problem ordinarily involve quite different arrangements of the stimulus-objects. In a discrimination between horizontal and vertical stripes, for example, the simultaneous problem presents both a horizontally and a vertically striped card on each trial, while the successive problem presents either two horizontally or two vertically striped cards on each trial. If, however, *Ss* are required to learn two discriminations *concurrently, e.g.* between horizontal and vertical stripes, and between large and small circles, the corresponding simultaneous and successive problems (really double-problems) can be so designed as to present identical stimulating conditions. As Table I illustrates, the two concurrent discriminations can be *segregated* in such a manner that the simultaneous and successive arrangements differ in the ordinary way, or they can be so *conjoined* that the stimultaneous and successive arrangements are identical and the two types of problem differ only in the direction of the rewarded response to two of the four pairs of stimuli. It should be clear, therefore, that the distinction between simultaneous and successive discrimination from which their designations derive is not an essential one. There remains another, more fundamental difference—in the perceptual organization underlying the solution of the two types of problem—which may perhaps best be analyzed when variation in stimulating conditions is controlled.

* Accepted for publication March 19, 1954.

[1] Phillip Weise and M. E. Bitterman, Response-selection in discriminative learning, *Psychol. Rev.*, 58, 1951, 185-195; M. E. Bitterman and Jerome Wodinsky, Simultaneous and successive discrimination, *Psychol. Rev.*, 60, 1953, 371-376; M. E. Bitterman and J. V. McConnell, The rôle of set in successive discrimination, this JOURNAL, 67, 1954, 129-132; Jerome Wodinsky, M. A. Varley, and M. E. Bitterman, Situational determinants of the relative difficulty of simultaneous and successive discrimination, *J. Compar. & Physiol. Psychol.*, 47, 1954, 337-340; E. F. MacCaslin, Successive and simultaneous discrimination as a function of stimulus-similarity, this JOURNAL, 67, 1954, 308-314. The two types of problem were named by W. S. Hunter, who was the first to suggest that their relative difficulty be studied (*J. Anim. Behav.*, 4, 1914, 215-222).

Previous research on the relative difficulty of simultaneous and successive discrimination has been concerned with the influence of four major variables. Two of these variables are *contact* and *contiguity,* whose effects (for historical reasons) have not yet been separated systematically. The most comprehensive study of the two variables, designed to show only their joint influence on the relative difficulty of simultaneous and successive problems, was made in a four-window jumping apparatus.[2] Under one set of conditions—designated as *component* conditions because they were thought to enhance the distinctiveness of the two members of each pair of stimuli—*S* jumped directly to one of the discriminanda (contact) which were situated in the two end-windows of the apparatus (noncontiguity), while the two center-windows contained identical gray cards. Under *configurational* conditions—so called because they were designed to retard the functional isolation of the discriminanda and promote the perception of some property of the pair as a whole—the stimuli to be discriminated were placed in the two center-windows (contiguity) and *S* jumped to one of the gray cards which were situated in the end-windows (thus making no contact with the discriminanda). The results of this experiment were quite unambiguous. Under component conditions the simultaneous problem was significantly easier than the successive, while under configurational conditions the successive problem was easier than the simultaneous. It is interesting to note that the conjoined double-problem provides yet a third method (independent of spatial separation and direct contact) for promoting the functional isolation of the discriminanda presented on each trial. Since, as may be seen in Table I, the characteristics common to horizontal and vertical stripes are quite different from those common to the large and small circles, the distinctiveness of the two members of each pair of stimuli presented to *S* is far greater in the conjoined than in the segregated problem.

Set is a third variable which has been studied. When each group of *S*s learns two consecutive problems of each type, marked differences in relative difficulty may appear with practice. Under component conditions (contact, noncontiguity) the absolute difficulty of the simultaneous discrimination does not change significantly from the first problem to the second, while the absolute difficulty of the successive discrimination decreases.[3] In two experiments with the two-window jumping apparatus (contact, contiguity) the gain in the successive group was so great that performance on the second simultaneous and successive problems did not differ significantly.[4] Under configurational conditions (noncontact, contiguity) the difficulty of the successive discrimination does not change markedly from the first problem to the second, but the decrease in the difficulty of the simultaneous discrimination is so great that the initial superiority of the successive group disappears entirely in the second problem.[5] These studies should be extended to include a much longer series of consecutive discriminations.

A fourth variable, whose influence may under certain conditions be very marked, is *stimulus-similarity.* MacCaslin, studying segregated simultaneous and successive double-problems in the two-window apparatus (contact, contiguity), found an in-

---

[2] Wodinsky, Varley, and Bitterman, *op. cit.,* 337-340.

[3] Wodinsky, Varley, and Bitterman, *op. cit.,* 337-340.

[4] M. E. Bitterman, A. D. Calvin, and C. B. Elam, Perceptual differentiation in the course of nondifferential reinforcement, *J. Compar. & Physiol. Psychol.,* 46, 1953, 393-397; Bitterman and McConnell, *op. cit.,* 129-132.

[5] Wodinsky, Varley, and Bitterman, *op. cit.,* 337-340.

crease in the relative difficulty of the successive discrimination as the similarity of the two members of each pair of stimuli increased.[6] These results were anticipated on the basis of the opportunity afforded by the segregated simultaneous problem for direct comparison of the two points on each afferent continuum. Saldanha and Bitterman, in a comparison of segregated and conjoined simultaneous double-problems, had previously demonstrated that opportunity for direct comparison may be critical for the discrimination of very similar stimuli.[7] Opportunity for comparison is, of course, controlled when conjoined double-problems are used in studies of simultaneous and successive discrimination.

It is reasonable that the analysis of discriminative learning should begin with component conditions, since they have been most widely studied. The solution of a simultaneous-component problem may be described as the development of a selective orientation ('approach') to one of the stimuli to be discriminated; here the two alternative responses available to *S* (right turn or left turn) are considered to be functionally equivalent, muscular discrepancies being ignored in favor of an emphasis upon the common adient characteristic.[8] Another possibility is that *S* learns to make one of two spatially opposed responses (regarded as distinct) to each arrangement of the stimuli to be discriminated, the effective cue being some unitary property of the arrangement as a whole ('configuration').[9] That this distinction is more than verbal has been demonstrated by Nissen, who devised an experimental situation for which contrary predictions are yielded by the two interpretations.[10] Nissen's results suggest that the concept of approach (or afferent selection) is better suited than the concept of response-selection to the solution by chimpanzees of a simple simultaneous problem (contact, contiguity).

In the successive-component problem the correctness of a given stimulus-card is contingent upon its position in space, and the behavior of *S* may be described in terms of 'approach' to certain card-position compounds (*e.g.* black-left or white-right). From this point of view the greater difficulty of the successive as compared with the simultaneous problem (under component conditions) is reasonable enough, but a number of disturbing questions remain. First, simple successive problems often are mastered so rapidly (even by naïve *Ss*) that an interpretation in terms of contingency or conditionality is not readily acceptable. At the very least, a supporting assumption about the relative simplicity of visual-kinesthetic compounding as compared, for example, with visual-visual compounding is required.[11] Furthermore, the

---

[6] MacCaslin, *op. cit.*, 308-314.

[7] E. L. Saldanha and M. E. Bitterman, Relational learning in the rat, this JOURNAL, 64, 1951, 37-53.

[8] K. W. Spence, The nature of discrimination learning in animals, *Psychol. Rev.*, 43, 1936, 427-449; H. W. Nissen, Description of the learned response in discrimination behavior, *Psychol. Rev.*, 57, 1950, 121-131.

[9] Harold Gulliksen and D. L. Wolfle, A theory of learning and transfer, *Psychometrika*, 3, 1938, 225-251; D. C. Teas and M. E. Bitterman, Perceptual organization in the rat, *Psychol. Rev.*, 59, 1952, 130-140; J. R. Turbeville, A. D. Calvin, and M. E. Bitterman, Relational and configurational learning in the rat, this JOURNAL, 65, 1952, 424-433.

[10] Nissen, *op. cit.*, 121-131.

[11] Nissen, Further comment on approach-avoidance as categories of response, *Psychol. Rev.*, 59, 1952, 161-167. To account in terms of an approach formulation for the results of Teas and Bitterman (*op. cit.*, 130-140) with the so-called 'two-situational' problem, it is necessary to assume that generalized visual-spatial conditionalities develop more readily than do simple component solutions involving the same visual stimuli.

rapid decline in the relative difficulty of the successive problem—so rapid in some situations that the difference in difficulty disappears entirely after only a single preceding problem—must be explained. Does the rat so readily develop a general ability to respond to contingency? The more rapid decline in the relative difficulty of the successive-component problem in the two-window as compared with the four-window jumping apparatus is especially troublesome from the point of view of compounding.[12] The greater spatial separation of the compounds in the four-window apparatus should increase their distinctiveness and facilitate learning. Finally, there is the question posed by the finding that segregated and conjoined successive presentations are not always functionally equivalent although they involve identical card-position compounds.[13]

An alternative to a consistent interpretation of discriminative behavior in terms of approach is a dual-process formulation which deals with the simultaneous solution in terms of approach and the successive solution in terms of response-selection. In a muscular sense, of course, the difference between the two alternative responses is just as great in the simultaneous problem as in the successive problem, but the functional importance of that difference may vary with the set of *S*. In any given situation the perception of *S* may be organized in terms of the difference between objects competing for a perceptually unitary response, or in terms of the difference between competing responses which tend to be elicited by some unitary feature of the environment. From this point of view, the greater initial difficulty of the successive as compared with the simultaneous problem under component conditions may be traced to situational determinants (contact and spatial separation) which emphasize the choice of stimuli rather than the choice of responses. Once a configurational set has been established, however, the solution of subsequent successive problems is facilitated; and the greater practice effect which is found in the two-window as compared with the four-window apparatus follows from the assumption that spatial separation impairs the stability of a configurational set. From this point of view, no equivalence of segregated and conjoined successive double-problems need be expected, since the configurational properties of the two sets of stimulating conditions may be quite different.

When the results of training under configurational conditions are taken into account some further support for a pure 'approach' formulation appears. A not inconsiderable advantage of this formulation is that it maintains a qualitative continuity among the four principal classes of problem, which are ordered quantitatively on the basis of the number of elements relevant to the solution of each. Thus, the simultaneous-component problem may be solved in terms of one element (*e.g.* approach horizontal stripes), the successive-component problem in terms of two elements (*e.g.* approach horizontal stripes in the right end-window), the successive-configurational problem in terms of three elements (*e.g.* approach gray in the right end-window when horizontal stripes are present), and the simultaneous-configurational problem in terms of four elements (*e.g.* approach gray in the right end-window when hori-

---

[12] Results for the two-window apparatus were reported by Bitterman and McConnell (*op. cit.*, 129-132) and for the four-window apparatus by Wodinsky, Varley, and Bitterman (*op. cit.*, 337-340).

[13] Wodinsky and Bitterman, Compound and configuration in successive discrimination, this JOURNAL, 65, 1952, 563-572.

zontal stripes appear in the right center-window). The order of initial difficulty obtained in a recent comparison of the four types of problem is exactly that which is suggested by an analysis of the number of relevant components, but again certain disturbing questions arise.[14] First, if the simultaneous-configurational solution requires the appreciation of a four-element or two-compound contingency, the readiness with which it is achieved by the rat—especially by completely naïve ones—is remarkable. Furthermore, there is no obvious explanation of the fact that in the second of two consecutive configurational problems the simultaneous discrimination is mastered as rapidly as the successive.

In terms of the dual-process formulation, contiguity and nonapproach may be said to promote configurational organization (thus facilitating the solution of successive problems) and to retard component organization (thus interfering with the solution of simultaneous problems). The rapid change with practice in the relative difficulty of the two problems is explained on the assumption that, in the course of the initial simultaneous-configurational problem, *S* develops a selective orientation to the discriminanda (component set) which carries over to subsequent simultaneous problems. (It is conceivable, of course, that even prolonged training would not reduce the difficulty of the simultaneous-configurational problem to the simultaneous-component level. An experiment involving a long series of consecutive problems would make it possible to determine whether there is an irrevocable functional difference between direct approach to a stimulus and approach to a neutral region adjacent to it—that is, whether the configurational conditions bring some new process permanently into play or merely hinder temporarily the development of a set equivalent to that which appears in simultaneous-component problems.) The only real challenge to the dual-process formulation which is offered by existing data lies in the fact that the initial difficulty of the successive-configurational discrimination is greater than that of the successive-component discrimination, and that this difference persists in the second of two consecutive problems.[15] On the assumption that successive problems are solved on the basis of a configurational organization, learning should be facilitated under conditions of contiguity and nonapproach, while the component conditions should have a retarding effect. It is possible, however, that direct approach makes the relevance of the discriminanda more apparent, and that this difference more than compensates for the greater difficulty of configurational organization under component conditions.

Problem

Further information about the processes responsible for simultaneous and successive discrimination was sought in this experiment. In the first stage, conjoined double-problems were studied under both component and configurational conditions. In the second stage, the relation between conjoined and segregated problems was studied; each *S* was trained on the segregated problem corresponding to the conjoined problem of the first stage. In the third stage, the relation between component and configurational conditions

[14] Wodinsky, Varley, and Bitterman, *op. cit.*, 337-340.
[15] *Idem.*

was studied by shifting each S trained on a component problem to the corresponding configurational problem, and vice versa.

Method

Subjects. The Ss were 48 experimentally naïve Albino rats, bred in the laboratory. They ranged in age between 3 and 4 mo. at the beginning of the experiment.

Apparatus. The apparatus employed was a four-window jumping stand. The circular jumping platform was equidistant from each of the windows (5.5 in. square) which were cut into the panels of a hemi-octagonal surround. Behind the windows was a feeding platform, and below them was a cloth net into which S fell after each incorrect response. The entire apparatus was painted flat black. The stimulus-cards were of five different kinds: (*a*) homogeneous mid-gray; (*b*) black-and-white vertical stripes (width of each stripe = 0.5 in.); (*c*) black-and-white horizontal stripes (equal in width to the vertical stripes); (*d*) small black circle (diameter = 1 in.) centered on a white ground; and (*e*) large black circle (diameter = 4 in.) centered on a white ground.

Preliminary training. Throughout the experiment the Ss were maintained on a 24-hr. feeding schedule. After adjustment to handling, they were fed wet mash on the feeding platform and then taught to jump gradually increasing distances (maximum = 9 in.) to open windows at the extreme right and extreme left only. The two central windows were open, but the Ss were not trained to jump to them, nor were they approached spontaneously. Manual guidance was used to provide equal experience with both end-windows. At the conclusion of the preliminary training, the Ss were divided into four equal groups matched for adjustment to the apparatus.

Stage 1. In the first stage of the experiment, each S was trained on a *conjoined double-problem* requiring a discrimination between horizontal (H) and vertical (V) stripes, and between large (L) and small (M) circles. Two of the four groups learned simultaneous problems and the other two groups learned successive problems (illustrated in a general way in Table I). For the *simultaneous-component* group, the center-windows contained mid-gray cards (G), and S jumped directly to the discriminanda which appeared in the end-windows, *i.e.* $HGGM$, $MGGH$, $VGGL$ and $LGGV$. Three of the Ss were rewarded for jumping to H and L, three to H and M, three to V and L, and three to V and M. Exactly the same stimulating conditions were encountered by the *successive-component* group, but the pattern of rewarded responses differed. For example, 3 Ss were rewarded for jumping to the right end-window in the situations $HGGM$ and $MGGH$, and for jumping to the left end-window in situations $VGGL$ and $LGGV$. Similarly, 3 Ss jumped left to $HGGM$ and $MGGH$, right to $VGGL$ and $LGGV$; 3 Ss jumped right to $HGGL$ and $LGGH$, left to $VGGM$ and $MGGV$; and the remaining 3 Ss jumped left to $HGGL$ and $LGGH$, right to $VGGM$ and $MGGV$. The treatment of the *simultaneous-configurational* group corresponded exactly to that of the simultaneous-component group, except that the discriminanda were in the center-windows and S jumped to the gray cards situated in the end-windows. Thus, for each S in the simultaneous-component group which was rewarded for jumping to the right end-window in the situation $HGGM$, there was an S in the simultaneous-configurational group which was rewarded for jumping to the right end-window in the situation $GHMG$, and so forth. In the same manner, the training of the *successive-configurational* group corresponded to that

of the successive-component group. The stimulating conditions encountered by the two configurational groups (simultaneous and successive) were, of course, identical, the two problems differing only in the pattern of rewarded responses.

Each *S* was given 10 trials per day by the method of correction. Each trial consisted of a series of one or more jumps to a given arrangement of the cards, terminating with a correct response. A maximum of three free jumps was permitted

TABLE I

ILLUSTRATION OF SIMULTANEOUS AND SUCCESSIVE DOUBLE-PROBLEMS, SEGREGATED AND CONJOINED

L=large circle; *M*=small circle; *H*=horizontal stripes; *V*=vertical stripes. The spatial arrangement of the symbols shows the spatial disposition of the stimulus-cards, *e.g.* *LM*=large circle to the left and small circle to the right.

	Simultaneous		Successive	
	Stimulus-arrangement	Rewarded response	Stimulus-arrangement	Rewarded response
Segregated	*LM*	right	*MM*	right
	ML	left	*LL*	left
	VH	left	*VV*	left
	HV	right	*HH*	right
Conjoined	*HM*	right	*HM*	right
	MH	left	*MH*	right
	VL	left	*VL*	left
	LV	right	*LV*	left

on each trial, and after three consecutive errors *S* was manually guided in the correct direction. The criterion of learning was one errorless day, and as each *S* met the criterion it was shifted (on the following day) to the problem of the second stage.

Stage 2. In this stage of the experiment, each *S* was trained (in the same manner and to the same criterion as in Stage 1) on the segregated double-problem corresponding to the conjoined problem mastered in the first stage. The relation between the two types of problem is shown in Table I.

Stage 3. In this stage of the experiment, each *S* was trained (in the same manner as before) on the conjoined problem of the first stage, except that *S*s trained under configurational conditions in the first and second stages were trained under component conditions in the third stage, while *S*s previously trained under component conditions were trained under configurational conditions in the third stage.

RESULTS

The course of learning in the first stage of the experiment is plotted (in terms of mean initial errors) in Fig. 1. The performance of the four groups is summarized quantitatively in terms of mean initial errors and mean total errors (initial plus repetitive) in Table II. Table II also shows comparable scores made by four groups of a previous experiment on corresponding problems in which the two discriminations were segregated and consecutive

rather than conjoined.[16] In both experiments learning was most rapid on the simultaneous-component problem, and least rapid on the simultaneous-configurational problem, but in the present experiment, unlike the previous one, there was no significant difference between the two successive problems.[17] Under component conditions, each of the conjoined problems of the present experiment was equal in difficulty to the corresponding segregated

TABLE II

PERFORMANCE ON THE FOUR CONJOINED DOUBLE-PROBLEMS OF STAGE 1 AND PERFORMANCE ON THE CORRESPONDING SEGREGATED PROBLEMS OF A PREVIOUS EXPERIMENT

	Initial errors				Total errors			
	Stage 1	Diff.	Previous	Diff.	Stage 1	Diff.	Previous	Diff.
Simultaneous-	24.3		24.8		37.4		38.0	
component		25.4*		37.5*		40.1*		49.3*
Successive-	49.7		61.8		77.5		87.3	
component		1.3		41.2*		−0.2		62.4*
Successive-	51.0		103.0		77.3		149.7	
configurational		49.4*		55.2*		71.7*		75.0*
Simultaneous-	100.4		158.2		149.0		224.7	
configurational								

* Significant beyond the 1-% level by Wilcoxon's nonparametric test for unpaired replicates.

problem of the earlier experiment. Under configurational conditions, however, the conjoined problems were considerably easier than the segregated problems of the earlier experiment.

All groups showed a great deal of transfer from the conjoined problems of Stage 1 to the segregated problems of Stage 2, although all were adversely affected by change in the arrangement of stimulus-cards. As Table III shows, each configurational group made significantly more errors in Stage 2 than did the corresponding component group.

In the third stage of the experiment, the order of difficulty of the four conjoined problems was exactly the same as that obtained in Stage 1. The simultaneous-component problem was easier than the successive-component problem, the successive-configurational problem was easier than the simultaneous-configurational problem, and there was no significant difference between the two successive problems. Since the data of Stage 3 summarized in Table IV are based only on 19 days of training, data for the first 19 days of training in Stage 1 are presented for purposes of comparison.

[16] Wodinsky, Varley, and Bitterman, *op. cit.*, 337-340.

[17] Tests of statistical significance were made by a nonparametric method (Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Co., Stamford, Conn., 1949, 1-16).

Measures of transfer (in terms of percent-saving) also are presented in Table IV. Transfer was greatest from the simultaneous-configurational problem to the simultaneous-component problem, and least from the latter to the former. It may be noted that animals trained on the simultaneous-

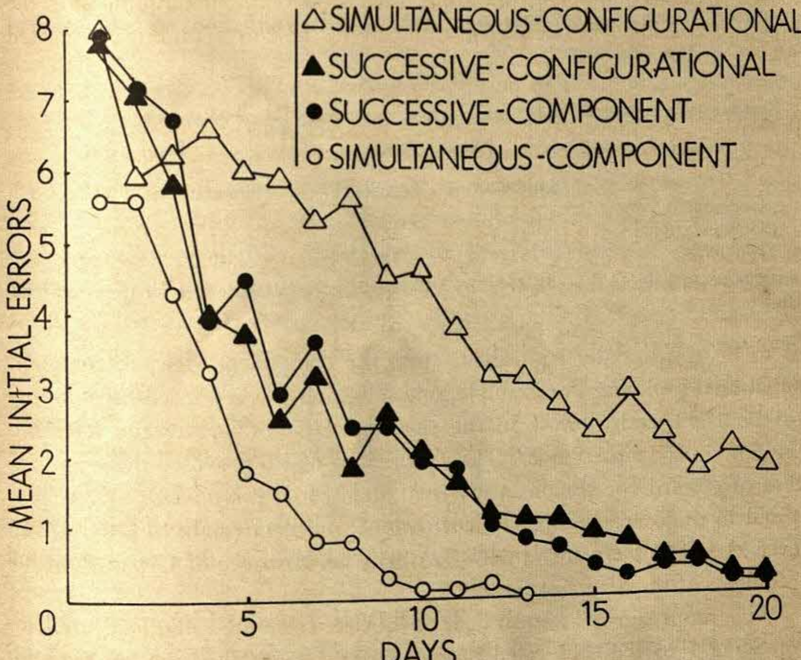


FIG. 1. THE COURSE OF LEARNING IN THE FOUR CONJOINED PROBLEMS OF STAGE 1

component problem in Stage 1 continued on the initial trials of Stage 3 to jump to the discriminanda, which had been shifted from the end-windows to the center-windows of the apparatus. Animals trained initially under configurational conditions showed no tendency in Stage 3 to jump to the gray cards.

DISCUSSION

In the simultaneous-component problem, the animal learns principally to approach or to avoid particular stimulus-cards. This conclusion is based on the fact that conjoined and segregated presentations are mastered at very much the same rate, that transfer from conjoined to segregated presentations is high, and that animals trained on a simultaneous-component problem tend to jump to the stimulus-cards encountered in a subsequent

configurational problem. Further evidence for this view was obtained by training a new group of 10 rats (Group *P*) on a two-situational version of the conjoined simultaneous-component problem. The procedure was in all respects comparable to that employed for the simultaneous-component group in Stage 1 of the main experiment, except that the problem involved only two of the four card-arrangements. The two situations were so selected

TABLE III

PERFORMANCE OF THE FOUR GROUPS ON THE SEGREGATED DOUBLE-PROBLEMS OF STAGE 2

	Initial errors		Total errors	
	Simultaneous	Successive	Simultaneous	Successive
Component	3.8	3.8	5.1	4.6
Configurational	15.3	21.3	21.3	29.6
Difference	11.5*	17.5*	16.2*	25.0*

* Significant beyond the 5-% level by Wilcoxon's nonparametric test for unpaired replicates.

as to be functionally equivalent, from the component view, to the four-situational problem. For example, one subgroup was rewarded for jumping to the right end-window in the situation *HGGM* and to the left end-window in the situation *VGGL*, a procedure which provided the same pattern of reward for responses to visual and spatial *components* as was provided in the corresponding four-situational problem (right to *HGGM* and *LGGV*, left to *VGGL* and *MGGH*). The performance of Group *P* on the

TABLE IV

RELATIVE DIFFICULTY OF THE FOUR CONJOINED DOUBLE-PROBLEMS IN STAGES 1 AND 3 (BASED ON PERFORMANCE DURING 19 DAYS OF TRAINING IN EACH STAGE) AND TRANSFER FROM THE FIRST TO THE THIRD STAGE

	Initial errors			Total errors		
	Stage 1	Stage 3	Saving (%)	Stage 1	Stage 3	Saving (%)
Simultaneous-component	24.3	4.2	82.7	37.4	6.5	82.6
Successive-component	49.4	20.0	59.5	77.0	25.4	67.0
Successive-configurational	48.4	16.4	66.1	74.0	24.1	67.4
Simultaneous-configurational	82.7	47.7	42.3	128.4	64.2	50.2

two-situational problem (29.9 initial errors and 45.9 total errors) did not differ significantly from the performance of the group trained on the four-situational simultaneous-component problem in Stage 1. Transverse or grouping effects do, however, appear in the simultaneous-component context.[18] Although transfer from the conjoined to the segregated presentation

[18] Bitterman, Spence on the problem of patterning, *Psychol. Rev.*, 60, 1953, 123-126.

of Stage 2 was high, it was not perfect, and transfer from the two-situational to the four-situational problem in the supplementary work with Group *P* was still less complete. After reaching criterion on the two-situational problem, the animals of Group *P* made an average of 6.4 initial errors and 9.4 total errors in the course of training to criterion on the corresponding (conjoined) four-situational simultaneous-component problem.

Solution of the simultaneous-configurational problem also appears to be based in considerable degree on the perception of components. This conclusion is suggested in part by the high transfer from the simultaneous-configurational to the simultaneous-component problem in Stage 3. It also is supported by the more rapid mastery of the conjoined as compared with the segregated simultaneous-configurational problem (on the assumption that the factor principally responsible for the difference is the greater dissimilarity, and consequently greater perceptual isolation, of the two stimulus-cards in each conjoined arrangement). There is little reason to believe that the simultaneous-configurational solution is based on a four-component conditionality; in all likelihood the animal learns merely to approach the gray card adjacent to each reinforced component. There is, however, a grouping effect of considerable magnitude in this problem, as is evidenced by the relatively large number of errors made by the simultaneous-configurational group transferred in Stage 2 from the conjoined to the segregated problem. This result probably is due in large part to the spatial contiguity of the discriminanda. It is clear, furthermore, that a component organization is not the organization most readily achieved under configurational conditions. A fresh group of 10 animals (Group *N*) subsequently trained to criterion on the two-situational problem corresponding to the four-situational simultaneous-configurational problem of Stage 1 (*e.g.* right to *GHMG* and left to *GVLG*) showed no positive transfer to the four-situational problem.

It is interesting to note that the two-situational problem of Group *N* had no more objective relation to the four-situational simultaneous problem than it had to the four-situational successive problem. It was, in fact, a *generalized configurational problem,* in that it comprised the two situations common to the simultaneous and successive conjoined double-problems. Thus, the situations right-to-*GHMG* and left-to-*GVLG* constitute a successive problem when combined with the situations right-to-*GMHG* and left-to-*GLVG,* while they constitute a simultaneous problem when combined with the situations left-to-*GMHG* and to right-to-*GLVG*. In the same manner, the problem of Group *P* was a *generalized component problem.* Functionally, the generalized component problem was almost entirely

equivalent to the four-situational *simultaneous* problem—the two problems were learned at about the same rate and transfer from the first to the second was very high. The generalized configurational problem was more closely related functionally to the four-situational *successive* problem. Although no transfer data were obtained, the learning scores of Group *N* (47.9 initial errors and 72.7 total errors) were quite similar to those of the successive-configurational group in Stage 1.

Although the successive-component and successive-configurational problems of Stage 1 were mastered at the same rate, transfer from each to the other was quite incomplete. It is necessary to conclude, therefore, that two discriminative problems involving identical discriminanda may be learned at the same rate by two processes which are at least partially distinct. That the successive-component solution may be based on approach to visual-spatial (card-position) conjunctions is suggested by the equivalence of the segregated and conjoined presentations which are identical in that respect. A strong configurational tendency in the solution of the successive-configurational problem is suggested by the great initial difference in difficulty of the conjoined and segregated arrangements, and by the considerable number of errors made in Stage 2 by animals transferred to segregated arrangements after conjoined training. The equivalence of the conjoined successive-configurational problem and the generalized configurational problem suggests, however, that the animals were responding in some degree at least to the presence of particular stimulus-cards rather than to certain global properties of the different card-arrangements. The transfer between the two successive problems (component and configurational) indicates a close relationship between response to a given card in a given position and a directional response in the presence of a given card.

It is interesting to note that all of the conjoined problems might conceivably have been mastered in terms of only one of the two discriminations—circles alone or stripes alone. In fact, however, the animals in all groups made both discriminations. This conclusion is based on the finding that the errors recorded in Stage 2 were for all groups randomly distributed over the four segregated card-arrangements.

Summary

Recent research on the factors affecting the relative difficulty of simultaneous and successive discrimination is reviewed, and an experiment is reported in which the two types of problem were studied under identical stimulating conditions. The results reveal a variety of discriminative processes, distinct yet interrelated, of which theoretical account must be taken.

ON MEMORY MODALITIES

By Hans Wallach and Emanuel Averbach
Swarthmore College

When one carefully observes how his memory contents come to him in recall, he is led to the conclusion that there are memory modalities just as there are sense modalities. Such observations can be discussed with psychologically unsophisticated people. When the question is raised with such a group how one recalls telephone numbers, there will be many who say they recall them "by the sound." A few will dissent; they recall them in script or print. Some will be uncertain or even incredulous when they hear these claims. Neither description seems to fit in their cases. They recall numbers as such, devoid of sound or shape. It is hard to say whether they use a conceptual memory which delivers numbers in abstract form for which verbal expression is sought only when the need arises (corresponding to the process by which one fits verbal expression to a thought, a necessity which leads to explicit searching when an original thought must be expressed) or whether they recall them in a verbal form that stems from memory associated with some motor center for speech.

We are not at all sure whether this exhausts the possibilities or characterizes correctly the non-sensory modalities used in the recall of numbers. It is also possible that in many people different modalities function in such close association with each other that a given content emerges at once in two modes. Yet, with regard to the existence of separate auditory and visual memory modalities we feel sure of our ground.

Different memory modalities may even harbor different products of learning related to the same performance. When the senior author adds numbers he ordinarily uses an abstract memory. New numbers, *i.e.* the results of individual additions, come to mind as entities that represent their arithmetical characteristics. When numbers are sub-vocally pronounced they are pronounced in English. If, however, long columns of figures have to be added, a shift to a different method will soon occur. Numbers are now read in German and the result of an addition is held in mind in German for the addition of the next number. His performance becomes effortless and seems to run off automatically like the recall or recitation of a well-memorized poem where the words are first recalled and then understood. He has had much practice with long addition in his youth in Germany. Associations then acquired apparently become available as the shift into German occurs. What interests us here is that the result of each succeeding addition is 'heard' or sub-vocally pronounced, he is not sure which. The process appears to be based entirely on verbal associations, whereas

* Accepted for publication May 20, 1954.

in the method now ordinarily employed numbers come to mind as mere numerical entities. The new method apparently evolved as German verbiage disappeared from all thought processes.

There is not much hope that the questions about the nature of different modalities of memory and their association can be answered by observations like those here reported. Clearly, experimental research is needed, but so far such observations have not influenced experimental work on memory. Yet, even if memory modalities were not interesting in their own right, there are good reasons why it is necessary to deal with them now. In the first place, certain experimental results can be interpreted when one assumes that multiple memory for a given content can occur when several modalities are involved. Secondly, we believe that the conception of memory modality is essential in the study of recognition. We begin by discussing this latter point.

It is a basic assumption in our thinking about simple recognition that a given perceptual experience arouses directly only a memory laid down in the corresponding memory modality, *i.e.* that memory modality which belongs with the same sensory modality which gave rise to the perceptual experience. With Köhler we believe that simple recognition is based on the similarity between the perceptual experience that gives rise to recognition and a more or less identical previous experience currently represented by a memory trace.[1] Such similarity can obviously manifest itself only if the memory trace is of the same modality as the perceptual experience. Thus, a novel verbal item that has been presented auditorily should not be recognized directly when the second presentation is a visual one. Only if the nonsense word is now 'read,' that is, if the visual experience is translated into some verbal mode, may the trace of the original auditory experience be aroused. It must similarly be assumed that a conceptual memory, *e.g.* the meaning of a word, can be aroused by, say, an auditory experience only indirectly; namely, through the arousal of the trace of a previous similar auditory experience and from there by way of a previously established connection to the conceptual memory item.

Results of a recent experiment on the effect of verbalization by Kurtz and Hovland can be interpreted in the light of this discussion.[2] Children were shown familiar objects and were asked to find and encircle the names of these objects on sheets on which they were listed and to pronounce them; or, in the case of the control group, to encircle the photographs on sheets that showed only photographs. A week later, tests of recall or recognition were unexpectedly given.

On the recall-test, *Ss* of the experimental group who had marked the names of the objects did better than the *Ss* of the control group who had marked the photographs. We agree with the authors that this is due to verbalization forced upon all experi-

[1] Hans Wallach, Some considerations concerning the relation between perception and cognition, *J. Personal.*, 18, 1949, 6-13.

[2] K. H. Kurtz and C. I. Hovland, The effect of verbalization during observation of stimulus-objects upon accuracy of recognition and recall, *J. Exper. Psychol.*, 45, 1953, 157-163.

mental *Ss* when they had to mark the names off the objects and had to pronounce them aloud, while the control *Ss* would verbalize only spontaneously and probably did so infrequently. We believe, however, that this verbalization effect depends on the formation of multiple traces for a given item. To an experimental *S* each item is, in effect, presented as a visual object and as a mimeographed word which is seen and pronounced. Thus, apart from the formation of conceptual memories which does not necessarily depend on verbalization, each presentation can give rise to two visual memories; namely, of the object and of the word, an auditory memory of the word pronounced, and perhaps a memory of still another modality left by the act of pronouncing. The last three would not arise in the control experiment unless *S* had spontaneously verbalized. There are a number of ways in which multiplication of traces for a given content could conceivably favor recall.

The results of the recognition-test require a more detailed discussion. In both groups recognition was tested for half the items with photographs and for the other half with mimeographed words. Thus, there were four different results arising from the four combinations of two learning- and two test-conditions. Visually learned material was tested with visual items or with verbal items, and visually *and* verbally learned material was tested with visual or verbal items. In three of the four combinations could direct recognition regularly operate in the test, because the items presented in the test-situation had occurred identically in the learning-situation. Only when merely visually learned material was tested with verbal items was this not the case. It is for this last combination that Kurtz and Hovland obtained by far the lowest mean recognition-score; namely, 5.46 as against 6.38, 6.58, and 6.71 for the other three combinations where direct recognition could operate.

Yet, on the average more than five correct recognitions out of a total of eight possibilities were made in this combination where presumably direct recognition would not take place. It must be assumed either that, with visual presentation, spontaneous verbalization was very frequent or, more likely, that indirect recognition played a major rôle in the test. Indirect recognition could take place in the following manner. *S* could not fail to be aware of the meanings of the familiar objects that were presented to him (although the task of finding and marking them on the sheet did not require it) and this awareness should leave a memory in some transsensory modality. When the verbal items were presented in the test, awareness of their meaning could lead to recognition in this mode. Whatever the nature of indirect recognition may be, it has equal possibilities in all four experimental combinations. That the combination of visual learning with verbal test gave a markedly lower score than the other three combinations should therefore be attributed to the failure to permit operation of direct recognition which singles out this combination.

Kurtz and Hovland gave a different interpretation to the results of the recognition-tests. They see in them a confirmation of the hypothesis that "the accuracy of retention would be increased by verbalization at the time of stimulus-observation."[3] This hypothesis predicts a difference between the score for the combination of visual learning with verbal test (5.46) and the score for combination of visual and verbal learning with verbal test (6.38), and this difference was obtained. It also predicts, however, a difference between the score for the combination of visual learning with

[3] *Op. cit.*, 162.

visual test (6.71) and the score for the combination of visual and verbal learning with visual test (6.58) in favor of the latter, and this difference was not obtained. The authors propose additional hypotheses to account for this finding. The largest of all differences—namely, the one between the score for the combination of visual learning with verbal test (5.46) and the score for the combination of visual learning with visual test (6.71)—cannot be predicted from a verbalization hypothesis alone. The outcome of the three comparisons is, of course, in agreement with our own approach which simply predicts a lower score for the combination of visual learning with verbal test, because it does not permit direct recognition to operate.

Our own experiment was an attempt to demonstrate the existence of memory modalities in a manner that involved some confirmation of our basic assumption; namely, that direct recognition requires the perceptual event which evokes the recognitive process and the trace which is aroused to be of the same modality.

Our interpretation of Kurtz and Hovland's results was based on two assumptions: (1) Memory modalities make possible the formation of multiple memory traces for a given content, and an ensuing duplication may favor retention or recall or both.[4] (2) Direct recognition can take place only if the test-item is perceptually similar to the event that left the pertinent trace, *i.e.* is of the same modality. The result of the present experiments can be understood from these two premises. The learning situation was such that the material was acquired in at least two modalities, and the test of recognition was varied in such a manner that either both modalities could come into play or only one. A higher recognition-score was expected in the former case. The particular form of the test situation depended on our notion of direct recognition.

Nonsense words were presented on a memory drum and *S* was asked to read them alternately forward and backward. After that *S* was unexpectedly tested for recognition of the words. The test-series was also presented on a memory drum, though at a faster rate, and *S* was asked to state promptly, before the next word appeared, whether the given word had occurred before or was new. Inasmuch as in the learning situation the words had to be read aloud, there was for *S* an occasion to form multiple traces for all words, visual traces from seeing the words on the drum, auditory traces from hearing them pronounced, and perhaps traces left by the act of pronouncing. In the test, the words were again presented visually and this made direct arousal of the visual traces possible. Beyond that, however, we expected *S*

[4] The term memory-trace is not meant to suggest a particular theoretical position. It is merely a convenient way to refer to the memory of an individual psychological event.

to 'read' the words, *i.e.* to translate the visual experience into some verbal mode. The resultant mental process could then arouse traces left by the reading of the words in the learning situation.

This is where the fact becomes important that in the learning situation half the words were read forward and half backward. We assume that *S* 'read' the words forward in the test (if he read them at all) and, therefore, that the mental process of reading could lead to the arousal of original reading traces only in the case of those words which originally had been read forward. On the verbal level, a word read forward and a word read backward are very different entities, and the experience of reading a word forward in the test is so different from the previous experience of reading it backward that it cannot be expected that the trace of the latter becomes aroused when the former occurs. Thus, only visual traces could serve in the recognition of words originally read backward, whereas multiple traces were available for the recognition of words read forward. We could, therefore, expect a higher recognition for the words that were originally read forward, provided that the assumption was correct that *S* 'read' the words only forward in the test-situation.

We base this assumption on the fact that the rate of exposure on the test was far too high for reading each word twice, once forward and once backward. Thus, only if *S* knew at the moment when each word appeared whether it had been read forward or backward in the learning situation could he read the words in the right direction. To know this, however, the visual trace of the word would have to be aroused first, because without such an arousal memory mediating this knowledge could not come into play. Yet, the arousal of the visual trace would be equivalent to recognition. Thus, recognition in some verbal mode achieved by this route should be unnecessary and should not influence the recognition score. To repeat, the expected result was a higher recognitive score for forward read words. If such a result were obtained it might be open, however, to a different interpretation. Since it is harder to read words backwards than it is to read them normally, it might be argued that conditions in the learning situation favored the words read forward and thereby caused the result.

The interpretation given above can be checked with the control *S*s who learned by reading *all* of the words backward. Our interpretation of the expected result for the main experiment depends on the fact that in the test *S*s who had learned by reading the words alternately backward and forward would not know which words had originally been read backward. This would not be true for *S*s of the control group who would have read all

words backward. In the test, they may attempt to read the words backward and have a relatively high score in recognition, since both visual traces and memories in the verbal mode would become available. If such a result were obtained the alternate interpretation would have to be rejected.

EXPERIMENT I

Procedure. Sixteen two-syllable nonsense words each consisting of six letters were presented to the subject one by one on a memory drum at a rate of a word every three seconds. Fifty-two *S*s, undergraduate students at Swarthmore and Bryn Mawr Colleges, participated. They were divided into three approximately equal groups. The *S*s of Group A (18) were asked to read the first word forward, the second word backward, and so on; the *S*s of Group B (17) had to do the same alternately forward and backward reading but started by reading the first word backward. The *S*s of Group C (17) were asked to read all words backward. The task was presented to the *S*s as an experiment in reading. None of them expected that a test would follow the reading.

In the recognition-test which followed the rate of presentation was stepped up so that each word was presented only for 1.5 sec. To prepare *S* for the high speed of the test, he was first presented with a short trial, which consisted of the first and the last word of the original list and of three novel words. Thereafter he was given the recognition-test proper. The list started out with two novel words followed by the remaining 14 words of the original list arranged in a new sequence. These two presentations were introduced by appropriate instructions which included a statement that the lists would contain words that had not appeared before. *S* was urged to respond quickly to every word as soon as it appeared. For some *S*s—two from each of Groups A and B and one from Group C—the speed was too high; they became flustered and stopped responding. These cases were dropped from the study and were not included among the *S*s reported as participating.

Result. The results for the 14 words which had appeared both in the learning list and in the main list are given in Table I. For Groups A and

TABLE I

MEAN NUMBER OF RECOGNITIONS FOR SEVEN WORDS

Recognition for	Group A	Group B	Group C: Words read backward by	
			Group A	Group B
Words read forward	4.71 (67%)	5.25 (75%)		
Words read backward	3.12 (45%)	3.00 (43%)	5.36 (77%)	4.88 (70%)

B separate scores are given for the words that were read forward in the learning situation and those that were read backward. Inasmuch as different words were read backward by *S*s of Group A and *S*s of Group B, the results for Group C, where all words had been read backward, are also divided into sets of seven words dependent on whether the particular words

had been read backward by Group A or by Group B. Thus all results in the table are given as mean numbers of recognition of seven words.

The difference between the mean scores of words read forward and backward is reliable in the case of each group (Groups A and B) at the 0.01-level ($t = 4.18$ and 8.65). When the results for Groups A and B are combined mean recognition for words read forward amounts to 71% and for words read backward to 44%. This is the difference we expected.

The result for Group C, where *Ss* read all words backward, is to be compared with the scores for words read backward obtained with Groups A and B. Group C shows the higher amount of recognition and this difference is reliable in the case of each set of seven words at the 0.01-level ($t = 6.22$ and 4.48). In fact, recognition in Group C is as high as recognition for words read forward in Groups A and B, with the mean recognition for all words in Group C at 73%. This makes it possible to reject the alternate interpretation of the main result of the experiment which ascribes the lower recognition for words read backwards to the learning conditions.

Experiment I still has one essential shortcoming: We do not know how much genuine recognition the score of 44% obtained for the words read backward in Groups A and B actually represents. There are only two ways in which an *S* can respond in such a test and the chance level for a recognitive response is 50%. Thus, if we wish to know to what degree visual memory contributed to recognition, or whether it did so at all, we have to find out how frequently novel words introduced into the test-list are mistakenly recognized.

To do so, part of Experiment I was repeated and novel words were introduced in the test. Groups D and E of 15 *Ss* each followed the procedure employed for Groups A and B, except that the test was somewhat changed. It consisted here of five of the words read forward, five of the words read backward, and five novel words arranged in random order. Only 10 words from the learning-list instead of 14 were included in the test-list in order to make the latter comparable in length to the test-list of 16 words which had been given to Groups A and B.

The combined results for Groups D and E showed a mean recognition score of 78% for words read forward and of 56% for words read backward. The number of judgments indicating mistaken recognition of novel words amounted to 25%. The difference between recognition of words read forward and words read backward is reliable at the 0.001-level, as is the difference between recognition of words read backward and mistaken recognition of novel words. This latter difference, that is, the extent to

which recognition of words read backward exceeds the score for mistaken recognition of novel words represents the genuine recognition of words read backward.

EXPERIMENT II

This experiment was meant to serve a double purpose. In the first place we wished to repeat the essential parts of Experiment I with different word-material. In the second place we wished to explore another way of using the technique of alternately reading forward and backward in making part of the memory of a given item unavailable for recognition. Some words read backward were presented in the test in reverse spelling and therefore could not be recognized visually. If they were read forward, however, they would verbally be the same as in the original presentation where they had been read backward. Thus recognition would be due to verbal rather than to visual memory.

Procedure. The reading list consisted of 22 nonsense words of five letters, one letter less than the words used in Experiment I. Again the first and last word of the list were used in the trial-test and did not appear in the test proper. Twelve words were used for a repetition of Experiment I and the remaining eight were used for experimentation with reverse spelling.

Two groups of 15 *Ss* each (Groups A and B) had to read the words alternately forward and backward, and a third group (Group C, also 15 *Ss*) read all words backward. Again, words read forward by Group A were read backward by Group B, and vice versa, except for the eight words devoted to experimentation with reverse spelling. They occupied slightly different positions on the lists that were read by Group A and by Group B, such that four of them were read forward by *both* groups and the remaining four were always read backward. The four words read forward were used only to equalize the total number of words available for forward and for backward reading, and their recognition was not included in any of the results. The four words read backward appeared in the test-list for Group A with normal spelling, but for Group B they were spelled in reverse. This made it possible to compare recognition based on verbal memory with recognition based on visual memory for the same individual words.

Results. The part of Experiment II that was devoted to a repetition of the procedure of Experiment I confirmed the results obtained there. The scores of Groups A and B combined yielded a level of recognition of 78% for words read forward and of 62% for words read backward. The difference was reliable at the 0.001-level. As in the previous experiment recognition by the *Ss* of Group C who had read *all* words backward was as high as recognition of words read forward; namely 76%. This score was also reliably different from the score for words that had been read backward by

the *S*s of Groups A and B. Thus all essential results of Experiment I were obtained again.

Recognition of the words read backward that were presented in the test spelled in reverse was as high as 60% and was not essentially different from recognition of the same words when they were tested in normal spelling (64%). Such scores should represent genuine recognition by a reliable margin. That there was an appreciable recognition of words read backward and tested in reverse confirms once more our contention that separately usable traces were left in our learning presentation.

Conclusions

Contents of memory are usually defined in terms of the external events to which they refer. It seems, however, to be more adequate to consider instead the psychological processes that occur when the external events are perceived. This becomes important when a given external event leads to a number of different psychological processes each of which may leave its own memory trace. For instance, when nonsense syllables are visually presented *S* will not only see each syllable, he will also silently pronounce it and two psychological processes will then result from one objective event, each of which may leave its trace. If instead of nonsense syllables, meaningful words are presented and *S* is also aware of the word's meanings, there may be for each item still another trace which is left by the process of being aware of the meaning. Evidence has been presented which shows that such multiple traces for a given external event enhance scores in tests of recall and recognition.

More important, however, than the influence of multiple traces as such is an issue which is related to the nature of recognition. In the absence of a set, recognition is based on the similarity between the perceptual process which gives rise to recognition and the memory of the pertinent previous experience. Here it becomes important that different traces pertaining to an external event stem from psychological processes which are of different modality and thus dissimilar to each other. No direct recognition should, therefore, occur between a process and a trace of different modality, although the two pertain to identical external events. Predictions derived from this premise were confirmed.

GESTALT LAWS OF ORGANIZATION AND ORGANISMIC THEORY OF PERCEPTION: EFFECT OF ASYMMETRY INDUCED BY THE FACTOR OF SIMILARITY ON THE POSITION OF THE APPARENT MEDIAN PLANE AND APPARENT HORIZON

By SEYMOUR WAPNER and HEINZ WERNER, Clark University

Gestalt psychology formulated a number of principles of organization that describe how stimuli are experienced in terms of figure-ground, part-whole relationships, and other properties of patterns. The efficacy of these principles has been demonstrated mainly in phenomenal terms by showing, for instance, that perceived patterns depend on such conditions as the degree of similarity or contiguity among the stimuli. The present study attempts to demonstrate that such autochthonous factors not only produce patterning of the perceived sensory material but also influence the perception in terms of stimulus-organism relationships.

The autochthonous factor introduced in the present study is that of similarity. The methodology, which will be discussed more fully below, consists in measuring displacements of the apparent median plane (straight ahead) caused by asymmetrical placement of objects in the field of view. We may illustrate this method by referring to an earlier experiment on the median plane. We distinguished there between two concepts of the median plane, viz. the objective median plane that is physically defined and the apparent median plane that is perceptually defined. The apparent median plane refers to that position in space which is experienced as straight ahead. In that experiment, a luminescent square was placed in the dark in such a way that its left edge, or its right edge, was fixated by the *S*s and placed in the objective median plane. It was found that this test-figure when extending asymmetrically to the left of fixation (right edge fixated) produced a shift of the apparent median plane relatively to the left in comparison with the situation when the test-figure extended to the right of fixation. In other words, the over-all effect of asymmetrical placement consists in a displacement of the apparent median plane toward the center of the figure. Thus, asymmetrical position of a test-figure causes a change in the space relationship between object and body which, it seems to us,

* Accepted for publication July 6, 1954. This investigation was supported by a research grant M-348 from the National Institute of Mental Health. We are indebted to V. K. Alexander, Monte M. Bliss, Patricia Fitzsimmons, and Emma Kraidman, who assisted in the research.

can be interpreted only in organismic terms. In our interpretation, given within the framework of sensory-tonic theory, we assume that asymmetric stimulation arouses a tendency in the organism so to change its state that a more harmonious (symmetric) stimulus-organism relationship ensues.[1]

As is well known, this may be achieved on a primitive level by the actual turning of the body toward the center of the figure (tropotaxis); if actual turning cannot be achieved, as is the case with our own *Ss*, we assume that a change takes place in the distribution of muscular-tonic activity in the body which reflects itself in a shift in the position of the apparent median plane. By position of apparent median plane is meant the physical location of stimuli experienced as straight ahead. This shift is in the direction of the center of the figure, or stating it in another way, the perceptual effect is that of seeing an asymmetrical figure more symmetrically related to the body.

Since asymmetrical stimulation has thus been found to influence distinctly and measurably object-body relationships we utilized, in the present study, asymmetry as a methodological device for demonstrating that autochthonous factors such as similarity produce organismic effects. This was accomplished by introducing asymmetry of the similarity among stimuli rather than in terms of asymmetrical placement of objects.

Asymmetry in terms of similarity has been studied in two spatial arrangements. In the first experiment the effect of a left-right asymmetry is measured in terms of displacement of the apparent median plane; in the second experiment an up-down asymmetry produces a displacement of the apparent horizon.

Experiment I: Effect of the Autochthonous Factor of Similarity on the Position of the Apparent Median Plane

Procedure. The test patterns used each consisted of three 20-cm. squares placed next to one another. Two adjacent squares were made similar in respect to two characteristics, viz. color (red, green) and direction of grating (horizontal stripes, vertical stripes). Thus, for instance, one pattern consisted of two adjacent squares composed of vertical red stripes, the third square composed of green horizontal stripes (see Fig. 1). The establishment of similarity in terms of two characteristics rather than one only was done to insure a strong effect of this organizational factor.

---

[1] Heinz Werner and Seymour Wapner, Toward a general theory of perception, *Psychol. Rev.*, 59, 1952, 324-338; H. Werner, S. Wapner, and J. H. Bruell, Experiments on sensory-tonic field theory of perception: VI. Effect of position of head eyas and of object on position of the apparent median plane, *J. Exper. Psychol.*, 46, 1953, 293-299; Wapner, Werner, Bruell, and A. G. Goldstein, Experiments on sensory-tonic field theory of perception: VII. Effect of asymmetrical extent and starting positions of figures on the visual apparent median plane, *J. Exper. Psychol.*, 46, 1953, 300-307.

We shall refer to the two adjacent squares as the dual pattern; we will name the particular total pattern by describing the color and grating of the dual pattern.

The luminous test-pattern was placed in a dark room, 200 cm. from *S*, and could be moved horizontally in small steps in the fronto-parallel plane on a track; the test-pattern was initially so placed that its center part was in the objective median plane, physically defined as a plane coinciding with the median sagittal section of the body. Thus, the test pattern extended symmetrically to either side of fixation, but was asymmetrical in terms of configurational arrangement.

S's task was to fixate the center part of the pattern and to instruct *E* so to move the pattern that the fixated part appeared sraight ahead of *S*.

The tests were performed in a dark room. *S* sat erect throughout all trials in a

TABLE I

EXPERIMENT I: EFFECT OF SIMILARITY FACTOR ON POSITION OF APPARENT MEDIAN PLANE
(Analysis of variance)

Source of variation	*df*	Mean square	*F*	*p*
Between all individuals:	31	650.88	—	—
sex	1	40.96	<1.00*	$>.05$
sequence	7	400.60	<1.00	$>.05$
Individuals within sex and sequence	23	753.57	26.84	$<.01$
Order	7	24.28	<1.00	$>.05$
Conditions:	7	172.62	6.15	$<.01$
side of dual pattern (S)	1	992.25	35.34†	$<.01$
color of dual pattern (C)	1	.81	<1.00	$>.05$
grating in dual pattern (G)	1	45.39	1.62	$>.05$
$S\times C$	1	84.64	3.01	$>.05$
$S\times G$	1	75.91	2.70	$>.05$
$C\times G$	1	7.77	<1.00	$>.05$
$S\times C\times G$	1	1.60	<1.00	$>.05$
Square uniqueness	42	28.29	1.01	$>.05$
Error	168	28.08		
Total	255			

* Tested against individuals within sex and sequence.
† Tested against pooled error ($S\times C+S\times G$), $F=12.36$, $p<.01$.

chair with side supports. His head was held in position by an adjustable head-rest; by this device head and trunk position were kept constant throughout the trials.

The three squares making up the luminous test-pattern were the front face of light boxes; the front face consisted of ground glass on which were pasted strips of tape that light was emitted between the strips, producing in darkness a pattern of horizontal or vertical colored lines. The emitted light was either green or red depending on the color of the bulb lighted in the box. Since it has been demonstrated that brightness gradients in test-patterns significantly affect the position of the apparent median plane, it was important to equate the brightness within the configuration.[2] This was accomplished by equating the red and green colors employed

[2] Duilio Giannitrapani, Effect of brightness gradients on the position of the apparent median plane, unpublished M.A. thesis, Clark University, 1953.

in the dark by means of flicker photometry. The total brightness of the test-pattern was low so that nothing was visible in the dark room except the pattern.

Eight color-grating patterns (see Table II below) were used. These total patterns are characterized by the color and grating of the dual pattern, as well as the placement of the dual pattern to the left or right of the third square. Thus, two values of each of three variables are represented in the eight test-conditions listed: side of dual pattern (left, right); color of dual pattern (red, green); and direction of grating in dual pattern (horizontal, vertical).

A combination of factorial and latin-square design was used. An 8×8 latin-square was chosen at random, and 2 men and 2 women were tested in each of the sequences making a total of 32 *S*s. The design permitted control of sequence and order, and also provided the possibility of separating the variance due to experimental conditions into single factors (Side (S), Color (C), and Direction (D) of dual pattern), double interactions ($S \times C$, $S \times D$, $C \times D$), and a triple interaction ($S \times C \times D$).

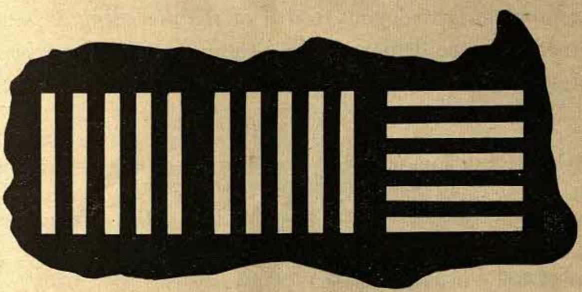


FIG. 1. SAMPLE OF TEST-PATTERN

The vertical white stripes were red or green; the horizontal, green. The black surround represents the dark room in which the test-pattern appears.

Measures employed. The position of the apparent median plane perceived by the *S* as straight ahead was measured in terms of deviation from the objective median plane. The objective median plane was determined by projecting the median sagittal plane of *S*'s head and body upon the fronto-parallel plane. Arbitrarily, positions to the left of the objective median plane were designated by minus values; to the right, by plus values.

Results. A summary of an analysis of variance is presented in Table I; means, mean differences, and *t*-tests are presented in Table II. Examination of the results show that the side of the dual pattern is the critical and significant factor. The position of the apparent median plane shifts relatively to the side of the dual pattern. The overall mean for dual pattern left is $+3.9$ cm. whereas for dual pattern right the overall mean is $+7.8$ cm. The mean difference is significant ($P < .01$).

The potency of the similarity factor is highlighted even more if one inspects the mean differences among the eight test conditions. There is

no overlapping of means for Conditions 1 through 4 (dual pattern left) compared to Conditions 5 through 8 (dual pattern right), and furthermore 14 of the 16 mean differences are significant.

There is no evidence of the efficacy of color (green vs. red) or direction of grating (horizontal vs. vertical) independently of the side of the dual pattern.

Experiment II: The Effect of the Autochthonous Factor of Similarity on the Position of the Apparent Horizon

Here the problem is analogous to that of the preceding experiment, viz. it is concerned with the demonstration of effects of similarity as measured in terms of spatial displacement; such displacement refers to the apparent

TABLE II

Experiment I: Effect of Similarity Factor on Position of Apparent Median Plane
(Means, mean differences, and *t*-tests)

	Test conditions				Mean differences						
No.	Side of dual pattern	Color of dual pattern	Grating in dual pattern	Means (in cm.)	2	3	4	5	6	7	8
1	left	red	horizontal	+4.7	0.4	1.4	1.5	1.9	3.3*	2.4	4.9†
2	left	red	vertical	+4.3		1.0	1.1	2.3	3.7†	2.8*	5.3†
3	left	green	horizontal	+3.3			0.1	3.3*	4.7†	3.8†	6.3†
4	left	green	vertical	+3.2				3.4*	4.8†	3.9†	6.4†
			Overall means=	+3.9							
5	right	red	horizontal	+6.6					1.4	0.5	3.0†
6	right	red	vertical	+8.0						0.9	1.6
7	right	green	horizontal	+7.1							2.5
8	right	green	vertical	+9.6							
			Overall means=	+7.8							

* Significant at or below 0.05 level; † Significant at or below 0.01 level.

horizon rather than the apparent median plane. We distinguish between two concepts of the horizon; namely, the objective horizon which is physically defined and the apparent horizon which is perceptually defined. The objective horizon is any line lying in the horizontal plane that passes through the eyes (with the body erect and the eyes gazing forward in a normal manner). The apparent horizon refers to that position in space experienced as being at eye-level.

It was found previously that, analogous to results for the apparent median plane, asymmetrical extension of a figure above or below the point of fixation affects the apparent horizon; that is, the apparent horizon shifts up when the figure extends up, and down when the figure extends down.[3] It was the purpose of this experiment to explore whether asym-

[3] K. M. Jaffe, Effect of asymmetrical position and directional dynamics of con-

metry introduced by the similarity factor has effects on the horizon in keeping with those found for the median plane in the experiment reported above.

Procedure. The tests were again performed in a dark room. *S* sat erect throughout all trials in a chair with side supports; his head was held in position by an adjustable head-rest and chin-rest; in this way the position of head and trunk was kept constant throughout all trials.

The same luminous color-grating patterns used in the previous experiment were employed here with the exception that they were rotated 90°. Again the dual pattern consisted of either the 'two upper squares' or the 'two lower squares' of the total pattern.

The light boxes were so mounted on a rack-and-gear device that it was possible to move the total pattern in small steps up and down along a plumb line at a distance of 200 cm. directly in front of *S*. The objective horizon was ascertained separately for each *S*; this was achieved by measuring the height of S's pupils above the floor. The center of the total pattern was always initially placed at the objective horizon to control for the effect of asymmetrical placement, a factor known to influence the position of the apparent horizon. In addition, brightness differences between the colors were ruled out by flicker photometry.

The eight test-patterns employed are listed in Table III. As in the previous experiment, there were two values for each of three variables represented in the eight test-conditions: position of dual pattern (above, below), color of dual pattern (red, green), and direction of grating in dual pattern (horizontal, vertical).

S's task was to fixate the center of the text-pattern and to instruct *E* so to move the pattern that its center appeared at eye level.

The experimental design was identical with that of the previous experiment: an 8 × 8 latin-square was chosen at random, and 2 men and 2 women were tested in each of the sequences, making a total of 32 *S*s.

The position of the apparent horizon was measured in terms of deviation from the objective horizon. Arbitrarily, positions above the objective horizon are designated by plus values; positions below the objective horizon are designated by minus values.

Results. A summary of the analysis of variance of Experiment II is presented in Table III; means, mean differences, and t-tests are presented in Table IV. Analogous to the results for the median plane, the position of the dual pattern is a critical and significant factor, *i.e.* the position of the apparent horizon shifts relatively up or down depending on whether the dual pattern is above or below. The overall mean for 'dual pattern up' is −7.6 cm., and for 'dual pattern down,' −12.6 cm. The mean difference is significant ($P < .01$).[4]

---

figurations on the visual perception of the horizon, unpublished M.A. thesis, Clark University, 1952.

[4] We may note here that, as has been found in a number of experiments, the apparent horizon under control conditions falls below the objective horizon. This explains why the results under all test-conditions are negative, *i.e.* below the objective horizon.

In addition to the main effect, we may consider the efficacy of color and direction of grating independent of the position of the dual pattern. There is no significant effect of color (red vs. green); however, direction of grating (vertical vs. horizontal) has a significant effect. The overall

TABLE III

EXPERIMENT II: EFFECT OF SIMILARITY FACTOR ON POSITION OF APPARENT HORIZON
(Analysis of variance)

Source of variation	df	Mean square	F	p
Between all individuals:	31	1,825.64	24.76	<.01
sex	1	3,647.40	2.07*	>.05
sequence	7	1,766.09	1.00*	>.05
Individuals within sex and sequence	23	1,764.56	24.93	<.01
Order	7	279.62	3.79	<.01
Conditions:	7	370.09	5.02	<.01
position of dual pattern (P)	1	1,571.62	21.32	<.01
color of dual pattern (C)	1	73.63	<1.00	>.05
grating in dual pattern (G)	1	384.89	5.22	<.05
P×C	1	206.47	2.80	>.05
P×G	1	30.19	<1.00	>.05
C×G	1	17.81	<1.00	>.05
P×C×G	1	306.02	4.15	<.05
Square uniqueness	42	105.92	1.44	>.05
Error	168	73.71	—	—
Total	255	305.78		

* Tested against individuals within sex or sequence.

mean for vertical grating in dual pattern is −8.9 cm., and for horizontal grating is −11.3 cm.[5]

SUMMARY AND CONCLUSIONS

Two experiments were carried out on the effect of asymmetry of a test-pattern induced by the autochthonous factor of similarity, on ego-centric localization in space. The asymmetrical arrangement was obtained by use of a test-pattern consisting of three adjacent squares; two adjacent squares were identical with respect to grating (horizontal or vertical stripes) and color (red or green) and differed in these two respects from the third square.

Though the three-square pattern was symmetrically placed with respect to the position of *S* (in the objective median plane or at the horizon, and fixated at the center), there was asymmetry induced by the factor of simi-

[5] The effect of direction of grating, not observed with the median plane, poses an interesting problem of configurational dynamics which requires further study.

larity. It was found that the apparent median plane or the apparent horizon was significantly displaced in the direction of the square which is similar to the center square.

Previous experiments have shown that asymmetrical object-stimulation induces a change in the organism that a more symmetrical relation obtains between the stimulus-object and the organism; this tendency we have termed "symmetrization."[6] In those experiments the asymmetry was achieved by so placing an oblong figure that it extended asymmetrically to the right or left of *S*'s objective median plane. Here, the asymmetry was not induced in terms of unequal extent but rather by asymmetrical relations among the parts of the figure.

TABLE IV

EXPERIMENT II: EFFECT OF SIMILARITY FACTOR ON POSITION OF APPARENT HORIZON
(Means, mean differences, and *t*-tests)

	Test-conditions				Mean differences						
No.	Position of dual pattern	Color of dual pattern	Grating in dual pattern	Means (in cm.)	2	3	4	5	6	7	8
1	up	red	horizontal	−10.2	5.8*	2.0	2.4	3.9	3.8	2.7	0.7
2	up	red	vertical	− 4.4		3.8	3.4	9.7†	9.6†	8.5†	5.1*
3	up	green	horizontal	− 8.2			0.4	5.9†	5.8*	4.7*	1.3
4	up	green	vertical	− 7.8				6.3†	6.2†	5.1*	1.7
			Overall means =	− 7.6							
5	down	red	horizontal	−14.1					0.1	1.2	4.6*
6	down	red	vertical	−14.0						1.1	4.5*
7	down	green	horizontal	−12.9							3.4
8	down	green	vertical	− 9.5							
			Overall means =	−12.6							

* Significant at or below 0.05 level; † Significant at or below 0.01 level.

The significance of this experiment lies in the demonstration that laws of organization, usually described in phenomenal terms, can be shown to have an effect (object-body symmetrization tendencies) that can be measured in behavioral terms.

[6] Werner and Wapner, *op. cit.*, p. 333.

SOME FACTORS INFLUENCING THE VARIABILITY OF VERNIER ADJUSTMENTS

By HERSCHEL LEIBOWITZ, University of Wisconsin

The ability to detect visually a discontinuity or 'break' in a contour, or to align two straight edges or bars, is known as vernier acuity. Because the variability of vernier settings is low (the lower the varibility the higher the acuity), the reading of vernier scales constitutes the visual task in the use of many optical instruments[1] Previous research has indicated that the standard deviation of such coincidence settings is 12 sec. of arc or less, and that this precision is unaffected by a thousand-fold change in luminance.[2] Baker determined the functional relation between vernier acuity and luminance and demonstrated that foveal vernier acuity, like many other measures of visual discrimination, is reduced systematically as the luminance level is lowered.[3] The aims of the present experiments are three in number: (1) To validate Baker's findings for the natural pupil. Her experiments were performed with the natural pupil; and the diameter of the pupil, which varies with luminance, is a factor in determining the properties of the retinal image;[4] (2) to determine whether the variability of the orientation with the test-object observed for grating acuity also exists in vernier discriminations;[5] and (3) to determine the effect of replacing the conventional dark bars viewed on a light background with light bars viewed on a dark background. In view of the reversal of the normal acuity-luminance relationship for parallel lines when light bars are substituted for dark, it is relevant to determine the effect of the reversal with the vernier.

PROCEDURE

Apparatus. The test-objects consisted of two rectangles the edges of which, at the beginning of a trial, were offset (Fig. 1A). *S*'s task was to adjust these rectangles that their edges appear to coincide (Fig. 1B). The rectangles, made of black paper, were mounted on glass with the lower limb stationary and the upper limb movable. The glass carrying the upper limb was mounted in a microscope stage which could be moved by *S* with the aid of a flexible cable and a micrometer head. The microme-

* Accepted for publication May 20, 1954. This research was supported in part under contract AF 18(600)-54 between the University of Wisconsin and the Psychology Branch, Aero Medical Laboratory, Directorate of Research, Wright Air Development Center, Air Research and Development Command, Wright Air Development Command, Wright Air Force Base, Ohio. The original report has been published as Wright Air Development Center Technical Report 53-200.

[1] D. H. Jacobs, *Fundamentals of Optical Engineering,* 1943, 86-88.

[2] Jacobs, *op. cit.,* 88; J. H. French, The unaided eye: III, *Trans. Opt. Soc.,* 21, 1919-1920, 127-156; G. L. Walls, Factors in human visual resolution, *J. Opt. Soc. Amer.,* 33, 1943, 487-505.

[3] K. E. Baker, Some variables influencing vernier acuity, *J. Opt. Soc. Amer.,* 39, 1949, 567-576.

[4] H. W. Leibowitz, The effect of pupil size on visual acuity for photometrically equated test fields at various levels of luminance, *J. Opt. Soc. Amer.,* 42, 1952, 416-422.

[5] Leibowitz, Some observations and theory on the variation of visual acuity with the orientation of the test-object, *J. Opt. Soc. Amer.,* 43, 1953, 902-905.

ter scale could be read to units of 0.002 mm. which corresponds to a visual angle of less than 0.2″ at 10 ft.

Light from a 300-w. bulb drawing 1.6 amp. A.C., passes through heat absorbing glass before illuminating a plate of ground glass located in back of the test-field. The test-objects were viewed in transmitted light with the natural pupil through a −1.0 diopter lens which requires emmetropic *S*s to accommodate for a simulated distance of 1 m. Constant current at the source was maintained with an ammeter

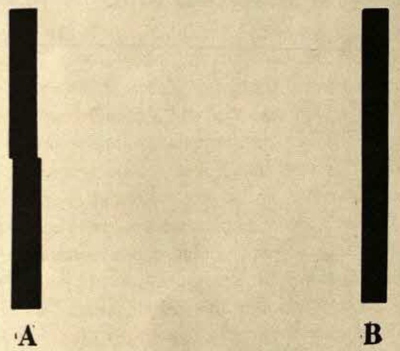


FIG. 1. VERNIER TEST-OBJECT
A, appearance at beginning of trial; B, appearance after *S* has set the components in alignment.

and rheostat. The maximal luminance was determined with a MacBeth illuminometer, and the luminance of the various test-fields employed computed from Wratten filters of neutral tint which were placed in the eyepiece of the apparatus.

Subjects. University undergraduates with normal acuity served as *S*s. They were instructed to align the rectangles that their adges appeared continuous. No time-limit would be placed on their adjustments. Each *S* was given several practice-trials before the experiments were begun. An individual session, including dark-adaptation, was completed within 1 hr.

For a given experimental session, *S* was first dark adapted for 10 min. and then light adapted to the lowest luminance of the test-field at which results could be obtained for him. *E* offset the test-object and *S* restored their coincidence by adjusting the upper limb. This procedure was repeated 20 times at each level of luminance. The direction and magnitude of offset was varied randomly, with equal frequency to right and left. *S*'s settings of 'confidence' were recorded by *E*. A single experimental session consisted of settings obtained at each luminance presented in an ascending order. The specific luminances employed and other details of the procedure will be described in the particular parts of the experiment where they are applicable.

RESULTS

To demonstrate the variability of the settings obtained, the results of one *S* (*W*) are shown in Fig. 2 for two levels of luminance: a low level (0.045 m.lam.) and one 10,000 times higher (450 m.lam.). These results—the frequency of *W*'s deviations from coincidence plotted at intervals of 3.3″—were obtained in 10 sessions

and with vertical light bars (30′ long and 6.5′ wide) against a dark background. They are typical of the results obtained from all the *Ss* in this study.

The distributions, as Fig. 2 shows, are approximately normal. The general effect of increasing luminance is to decrease the variability of *S*'s settings. At the lower luminance (0.045 m.lam.), the range is more than 1′ with 78% of the values falling within plus or minus 10″ of the true position of coincidence. At the highest level (450 m.lam.), *all* of the settings fall within the limits of plus or minus 10″. The *SD* of the distribution at the low level is 16.6″ while at the high level it is

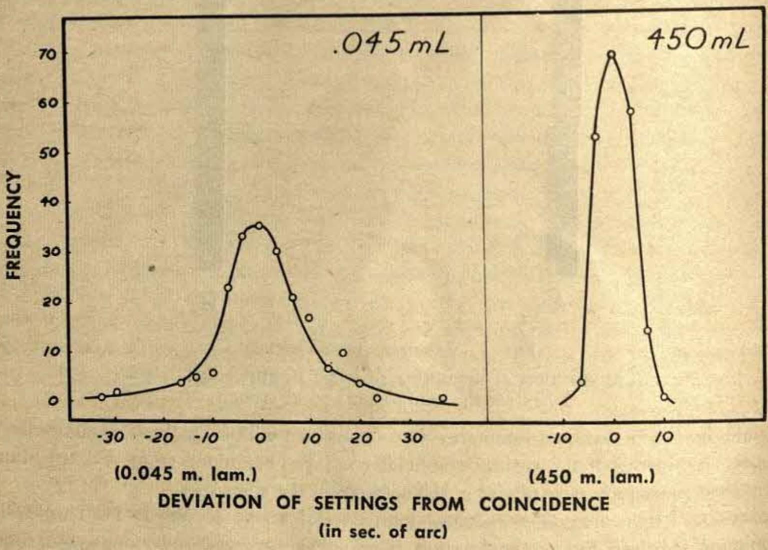


FIG. 2. FREQUENCY OF VERNIER ADJUSTMENTS AS A FUNCTION OF DEVIATIONS FROM TRUE COINCIDENCE

reduced to 5.8″. Statistical inspection of the mean settings obtained on the same or in different experimental sessions revealed no significant differences. The *SD* is taken as a measure of the variability of vernier adjustments and the results of the present experiment are reported in terms of this statistic.

Results were obtained in the present study under three conditions.

(1) In the first phase of the experiment, the test-objects consisted of two black rectangles (30′ long and 5′ wide) oriented vertically in a light field 2° square. Three *Ss* were used. Each served at three sessions, each session consisting of 20 coincidence adjustments at each of 10 luminances ranging from the lowest luminance at which the *S* could see the rectangles with foveal vision up to 200 m.lam. in steps of approximately 0.4 log units each. The results presented in Table I as the average of the standard deviations obtained for the three experimental sessions demonstrate the functional dependency of vernier variability upon luminance first described by Baker. As luminance is lowered, variability remains constant at first, then decreases slowly, and finally, at luminances near the lower limit of foveal vision, decreases rapidly. Variability at the lowest luminance is approximately 3 to 6 times

greater than at the highest levels. The increase in variability at low luminances exhibits considerable individual differences. *M* showed much less increase than the other *Ss*. It will also be observed that the lowest level (within 0.3 log units) at which results could be obtained is not the same for all *Ss*. The origin of these individual differences is not known, although the diameter of the pupil, which was not

TABLE I

VERNIER VARIABILITY FOR VERTICALLY ORIENTED DARK BARS

Log Luminance (m. lam.)	Ss M	R	W
−2.30	26.5	45.9	—
−2.00	12.4	37.4	29.7
−1.60	11.0	14.8	13.5
−1.30	13.0	12.3	9.3
−1.00	9.5	11.4	11.1
−0.50	10.2	7.6	10.6
0.10	11.0	6.9	9.5
0.70	8.4	8.7	5.9
1.10	8.4	7.7	7.6
1.40	8.4	6.5	6.9
2.00	9.9	6.8	8.1

controlled, may be a possible contributing factor. The data for all three *Ss* were averaged and plotted in Fig. 3 to demonstrate the increase in vernier variability at low luminance levels.

(2) In the second phase of the study, the same stimulus-components were presented in the horizontal position (H), 45° clockwise from vertical (R) and 45° counterclockwise from vertical (L). Rotation of the apparent position of the stimulus-components was achieved by means of a Dove prism in the eyepiece of the apparatus. Each *S* was given three experimental sessions at each orientation. The luminance-values of the test-field employed were selected from those used in the first

TABLE II

VERNIER VARIABILITY FOR VARIOUS ORIENTATIONS OF DARK BARS

Log luminance (m. lam.)	Ss											
	M				R				W			
	H	V	R	L	H	V	R	L	H	V	R	L
−2.30	19.3	26.5	35.3	41.4	—	45.9	—	—	—	—	—	—
−2.00	15.8	12.4	26.0	22.5	29.9	37.4	41.5	27.0	20.6	29.7	22.9	32.9
−1.60	13.2	11.0	15.1	12.5	28.0	14.8	25.4	20.9	15.6	13.5	17.2	16.5
−1.00	10.1	9.5	13.9	12.6	16.7	11.4	18.3	15.6	8.6	11.1	11.2	12.2
0.10	10.4	11.0	13.4	14.4	11.9	6.9	9.6	15.1	7.5	9.5	9.3	8.7
1.10	7.8	8.4	11.0	14.9	13.6	7.7	8.8	11.6	7.4	7.6	7.7	8.9
2.00	7.2	9.9	8.7	10.3	13.2	6.8	13.5	13.1	7.4	8.1	9.6	10.0

phase of the study so as to be of greatest utility in describing the variability-luminance function. The results for the H, R, and L positions, together with the appropriate data for the V position obtained from Table I, are presented in Table II. Inspection of these results indicates no consistent trends with respect to a given orientation position. If, however, the data for the average of horizontal and vertical positions (HV) are compared with the average for the oblique positions (RL), the superiority

of HV over RL occurs at all luminance levels for *M*, *W*, and at five of the six levels for *R*. Analysis of variance (see Table III) reveals this HV/RL superiority to be significant. The variability contributed by luminance is also significant as is that contributed by the *S*s. The data for all *S*s for the HV and for the RL positions have

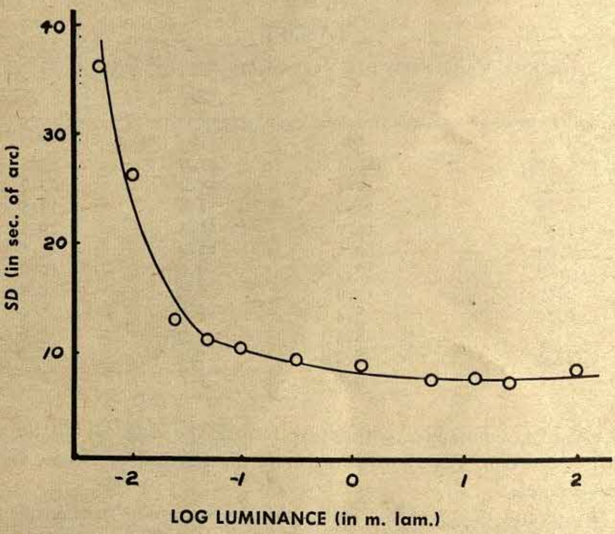


FIG. 3. STANDARD DEVIATION OF VERNIER ADJUSTMENTS FOR DARK BARS AS A FUNCTION OF LUMINANCE

been average and plotted in Fig. 4. This graph indicates that the effect of orientation is to displace the variability-luminance function along the variability axis, the oblique positions shifting the curve to higher variability values, the horizontal and vertical orientations to lower variability values.

(3) In the third phase of the experiment, the relationship between the test-object and the luminance of the background was reversed. Vertical light bars (30′ long and 6.5′ wide) were presented on a dark field. Results were obtained for *W*, an *S* who had served in the previous phases of the experiment, in 10 sessions at 6 luminances varying from 0.045 to 450 m.lam. A second *S* (*B*), who had not previously served in this study, was also given 10 sessions at luminances varying from 0.023 to 450 m.lam. The results of these *S*s, given in Table IV and Fig. 5, exhibit the same functional dependency of vernier variability as previously noted. The nature of the individual differences can be seen by reference to Fig. 5. The entire curve for *B* is about 5″ higher than for *W*, although both curves have the same shape. It will also be observed that both of these curves are located at higher luminance values than those obtained with dark bars on a light background. Apparently, the effect of various parameters, such as orientation, individual differences, and the test-object-luminance relationship, is to shift the function either laterally or vertically without changing its shape.

TABLE III

VARIANCE TABLE FOR ORIENTATION OF DARK BARS

Source	Sum of squares	df	Estimate of variance	F (Lum.×Ss×Orien. error term)	F (Lum.×Ss error term)
Luminance	4,577.1	5	915.4	$\frac{915.4}{16.5}=55.5$†	$\frac{915.4}{69.4}=13.2$‡
Ss	611.2	2	305.6	$\frac{305.6}{16.5}=18.5$‡	$\frac{305.6}{69.4}=4.40$*
Orientation	221.1	3	73.7	$\frac{73.7}{16.5}=4.47$*	
Lum.×Ss	694.3	10	69.4	$\frac{69.4}{16.5}=4.21$†	
Ss×Orien.	207.2	6	34.5	$\frac{34.5}{16.5}=2.09$	
Lum.×Orien.	236.3	15	15.8	$\frac{15.8}{16.5}$	
Lum.×Ss×Orien.	496.3	30	16.5		
Total	7,043.5	71			
Orientation (HV vs RL)	203.9	1	203.9	203.9/16.5=12.36†	

* Significant at 0.05 level. † Significant at 0.01 level. ‡ Significant at 0.001 level.

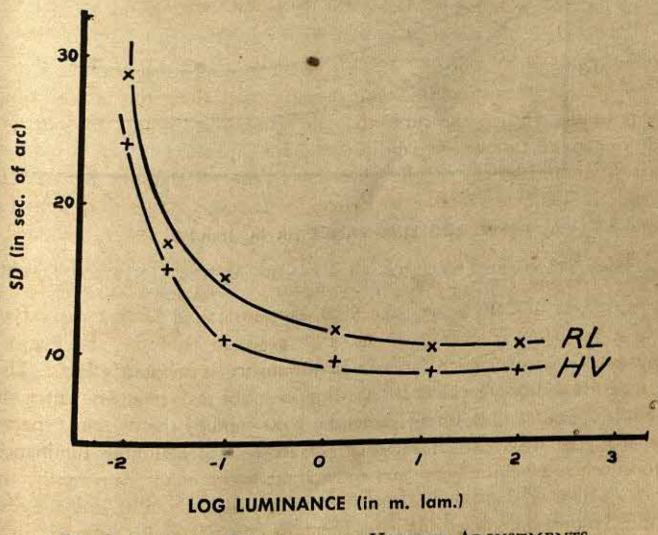


FIG. 4. STANDARD DEVIATION OF VERNIER ADJUSTMENTS
HV, dark bars oriented vertically and horizontally; RL, dark bars oriented 45° to the right and left of the vertical.

Discussion

The functional dependence of vernier variability upon luminance, as established by Baker, is confirmed by the present experiment for the natural pupil. This relationship is a familiar one in the field of visual discrimination. Discriminatory ability in general increases with luminance, rapidly at first and then more slowly, before

TABLE IV

Vernier Variability for Light Bars on a Dark Field

Log luminance (m. lam.)	Ss W	Ss B
−1.65	—	32.7
−1.35	14.5	19.0
−0.85	10.8	16.0
−0.35	8.6	13.3
0.65	7.0	10.3
1.65	5.8	10.8
2.65	5.7	10.8

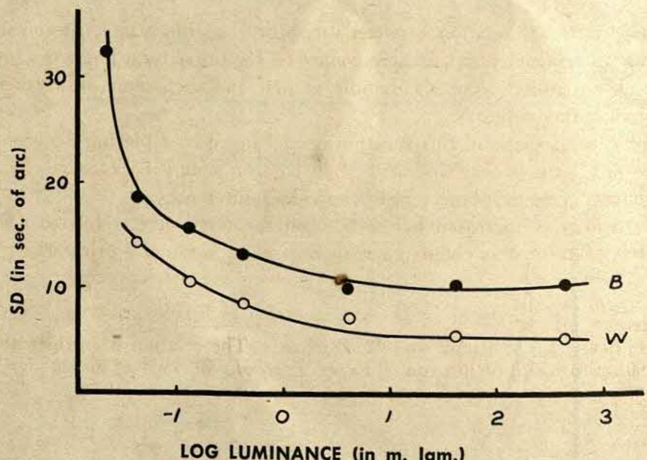


Fig. 5. Standard Deviation of Vernier Adjustments for Light Bars as a Function of Luminance
W, a practiced *S*; *B*, an unpracticed *S*.

reaching a level above which increase in luminance is without effect. This is true for grating, Landolt C, and single line test-objects, intensity discrimination,[6] stereoscopic acuity,[7] and has been shown by Baker and by the present experiment for vernier acuity as well.[8] The invariance of vernier acuity at high luminance values

[6] S. H. Bartley, The psychophysiology of vision, in S. S. Stevens (ed.) *Handbook of Experimental Psychology,* 1951, 921-984, esp. 958 f.

[7] C. G. Mueller and V. V. Lloyd, Stereoscopic acuity for various levels of illumination, *Proc. Nat. Acad. Sci.,* 34, 1948, 223-227.

[8] Baker, *op. cit.,* 570-575.

accounts for the report by French, who probably worked on the flat portion of the acuity-luminance function, that vernier acuity is not dependent upon luminance.[9]

The absence of any difference between the shape of the functions obtained using light and dark bars argues for the similarity of the visual task in the two situations. Apparently, as has been suggested by Jacobs, vernier adjustments consist of alignment of the *edges* of the test-objects.[10] This is consistent with the report of Berry, Riggs, and Richards who have shown that the width of the vernier test-object within the limits of 27 and 424 sec. of arc is an irrelevant variable.[11]

The superiority of the horizontal and vertical orientations over the oblique positions of the test-object is consistent with data obtained by other investigators who tested grating acuity. The magnitude of the effect is not large, the oblique positions increasing the variability about 20%. This difference is consistent, and is a factor where vernier adjustments or judgments are required in relation to obliquely oriented contours.

Summary

The functional relation between the variability of vernier adjustments and luminance was determined under various experimental conditions. The results are as follows:

(1) The functional relation between variability and luminance, previously determined with an artificial pupil, is here confirmed for the natural pupil. As luminance is increased variability decreases, rapidly at first and then more slowly before approaching a limiting value.

(2) The general shape of this function is unchanged by replacing the conventional dark bars on a light background with light bars on a dark background, by rotating the components of the test-object, or by individual differences.

(3) Variability is increased by 20% when the test-object is rotated 45° to the right or left of vertical as compared with vertical or horizontal orientation.

---

[9] French, *op. cit.*, 145.
[10] Jacobs, *op. cit.*, 87.
[11] R. N. Berry, L. A. Riggs, and W. Richards, The relation of vernier and depth discrimination to width of test rod, *J. Exper. Psychol.*, 40, 1950, 520-522.

THE DISTANCE-EFFECT IN THE TRANSPOSITION OF INTERMEDIATE SIZE BY CHILDREN

By HAROLD W. STEVENSON and M. E. BITTERMAN, University of Texas

After learning to discriminate between two points on some afferent continuum, the inarticulate *S* is more likely to transpose the solution when the testing stimuli are close to the training stimuli than when the two pairs are far apart. This *distance-effect,* which has appeared repeatedly in experiments with rat, monkey, chimpanzee, and preverbal child, is of considerable theoretical significance.[1] While it can be deduced from the nonrelational theory of Spence,[2] it contradicts a strict relational interpretation of learning which suggests that transposition should not be affected by a change in the absolute properties of the training stimuli as long as their relational properties are unaltered. In the light of recent evidence, which gives rise to serious doubts about the validity of an absolute theory, and which points to some sort of relational determination in the discriminative performance of inarticulate *Ss*,[3] it is important to note that the distance-effect is not entirely incompatible with a relational view. In the course of a relational solution *S* may learn something about the absolute properties of the training situation, and transfer of the relational solution may be impaired to the extent that those absolute properties are changed.

To account for the results of experiments on the discrimination of size by children in simple two-choice problems, a distinction between verbal (abstract) and preverbal (absolute) modes of learning has been suggested.[4] Children who were unable to verbalize the principle of solution showed the distance-effect, but older, verbal children did not. This difference was considered to support the hypothesis that the absolute "mechanisms assumed by Spence to underlie" transposition "in animals are also operative in the young child, and that with the development in older children of the capacity to employ verbal responses in such . . . situations, a shift occurs to the verbal type of control."[5] To be sure, the verbal ability of the

* Accepted for publication May 1, 1954.

[1] Harold Gulliksen, Studies of transfer of response: I. Relative versus absolute factors in the discrimination of size by the rat, *J. Genet. Psychol.*, 40, 1932, 37-51; Heinrich Klüver, *Behavior Mechanisms in Monkeys,* 1933, 1-387; K. W. Spence, Analysis of the formation of visual discrimination habits in the chimpanzee, *J. Compar. Psychol.*, 23, 1937, 77-100; M. R. Kuenne, Experimental investigation of the relation of language to transposition behavior in young children, *J. Exper. Psychol.*, 36, 1946, 471-490; Elizabeth Alberts and David Ehrenfreund, Transposition in children as a function of age, *J. Exper. Psychol.*, 41, 1951, 30-38.

[2] Spence, The differential response in animals to stimuli varying within a single dimension, *Psychol. Rev.*, 44, 1937, 430-444.

[3] E. L. Saldanha and M. E. Bitterman, Relational learning in the rat, this JOURNAL, 64, 1951, 37-53; C. B. Elam and Bitterman, The effect of an irrelevant relation on discriminative learning, *ibid.*, 66, 1953, 242-250; E. C. Wortz and Bitterman, On the effect of an irrelevant relation, *ibid.*, 66, 1953, 491-493; R. C. Gonzalez, G. V. Gentry, and Bitterman, Relational discrimination of intermediate size in the chimpanzee, *J. Compar. & Physiol. Psychol.*, 47, 1954, 385-388.

[4] Kuenne, *op. cit.*, 471-490; Alberts and Ehrenfreund, *op. cit.*, 30-38.

[5] Kuenne, *op. cit.*, 474.

older children together with their freedom from the distance-effect may be taken as evidence of an abstract level of functioning which is qualitatively distinct from the level displayed by younger children; there is, however, no need to conclude that the younger children (or even the infrahuman *Ss* with which they have been classed) function *solely* on an absolute basis, despite their lack of verbal ability and their dependence on distance. Instead, the somewhat general assumption may be made that, in the course of discriminative training, the younger child (or infrahuman *S*) may learn, not only the relation between positive and negative stimuli, but also something about the absolute properties of those stimuli (*e.g.* the region of the afferent continuum which they occupy), and the two sets of properties may be so linked that the stability of performance will be impaired in proportion to the extent of absolute change. From this point of view the difference in the behavior of older and younger children in previous experiments reflects, not a difference between an abstract, verbally-mediated relational process and a purely absolute one, but a difference between two relational processes—one that is abstract and one that is more closely bound up with the absolute properties of particular situations.

In the experiment to be reported, preschool children were trained to choose the intermediate member of a set of three stimulus-objects which differed in size. Then tests for transposition were made with new sets of objects either one or five steps removed from the training set. The appearance of a distance-effect in this experiment could be understood on the assumption that certain relative and absolute properties of the training stimuli were functionally linked in the original solution. The mere occurrence of transposition at the near point would indicate a relational process, since in the case of the intermediate-size problem the absolute theory predicts no transposition at all,[6] while a decline in transposition at the far point would indicate some dependence on absolute properties. The appearance of a distance-effect could not, on the other hand, be understood in terms of a theoretical dichotomy between verbal and absolute processes. With *Ss* capable of verbal solution, there should be as much transposition at the far point as at the near. With *Ss* incapable of verbal solution, there should be no transposition at either point.

Method and Procedure

Subjects. The *Ss* were 24 preschool children between the ages of 4 and 6 yr. Children of this age were selected because they can learn the intermediate-size problem with relative ease, although for the most part they are unable to verbalize the basis of solution.

Apparatus. The apparatus consisted of (*1*) a small, brightly colored 'Easter egg,' cut from paper, which served as the goal-object; (*2*) a series of eight blocks differing in size which served as the discriminanda; (*3*) a large tray on which the three blocks used on each trial could be arranged in a horizontal row; and (*4*) an opaque screen which could be interposed between *S* and the tray while the blocks were being arranged for each trial. The tray and screen were painted flat black. The eight blocks, designated in order of size by the numbers *1–8*, were unpainted

[6] Spence, The basis of solution by chimpanzees of the intermediate size problem, *J. Exper. Psychol.*, 31, 1942, 257-271; Gonzalez, Gentry, and Bitterman, *op. cit.*, 385-388.

squares cut from 1/4-in. Masonite. The area of the smallest block was 4 sq. in., and the area of each successive block increased by a factor of 1.4.

Procedure. *S* was seated before the tray on which the three blocks to be used in training were set in the first of a predetermined, random series of horizontal arrangements. *E* explained: "This is a game called 'Find the Easter Egg.' See the Easter egg? And see these blocks?" (pointing to the blocks). "Well, I'm going to hide the Easter egg under one of these blocks and I want to see if you can find it. I'll hide it under the same block every time. So let's see how often you can find it." *E* lowered the screen and hid the egg under the block of intermediate size. Then he raised the screen, exposing the blocks to *S*, and said: "You pick the one you think it's under." *S* was allowed only one choice on each trial. If the choice was incorrect, *E* lowered the screen and made some encouraging comment, such as "No, it wasn't under that one. But see if you can find it next time." After correct choices the procedure was the same except that *E* made some such comment as "Good: You found it. Now see if you can find it again." Training continued to a criterion of five consecutive correct choices.

Immediately after *S* had met the criterion of learning, a second *E* (who had been recording the data) attracted *S*'s attention and asked him to complete a simple drawing of a 'little boy' or a 'little girl.' *S* was occupied in this manner while the blocks used for training were removed from the tray and those to be used for the first test of transposition (T_1) were arranged. Approximately 2 min. after the last training trial, it was suggested to *S* that the 'game' be continued, and the first transposition series was begun. It consisted of six differentially reinforced trials with a set of stimuli either one (*near-test*) or five (*far-test*) steps removed from the training set. The intermediate stimulus in the transposition set was reinforced. Following these trials, *S* was asked to do further work on his drawing for a period of about 2 min. while the blocks employed for T_1 were replaced with those to be used in the second transposition test (T_2). T_2 was identical to T_1, except that if T_1 was a far-test T_2 was a near-test, and vice versa.

At the conclusion of T_2, *S*'s ability to verbalize the concept of *middleness* was explored. Indicating the set of blocks used for T_2, *E* asked *S* to identify the "middle-sized" or the "medium-sized" block (both terms were tried on the assumption that some *S*s might be more familiar with one than the other) and then, in sequence, the biggest, the middle-sized, the smallest, and the middle-sized blocks. Identical questions were asked about a set of blocks at the opposite end of the series. Finally, *S* was asked: "How did you know which one the egg would be under?" This question was rephrased in several ways to give *S* every opportunity to verbalize the basis of solution.

The design of the experiment is summarized in Table I. Of the 24 *S*s, 12 selected at random comprised Group I and the remaining 12 comprised Group II. Half of each group was trained with blocks *1-2-3,* while the remaining half was trained with blocks *6-7-8*. For Group I, T_1 (near-test) was made with a set of blocks one step removed from the training set—*2-3-4* in the case of *S*s trained with *1-2-3,* and *5-6-7* in the case of *S*s trained with *6-7-8;* T_2 (far-test) was made with a set of blocks five steps removed from the training set—*6-7-8* for *S*s trained with the small set and *1-2-3* for *S*s trained with the large set. For Group II, T_1 was the far-test and T_2 was the near-test.

Results

The training problem was learned very rapidly; the mean number of trials (including the five criterion trials) was 12.7, and the corresponding median was 11.0. The small number of trials may be explained in part by the relatively uncomplicated training conditions, and in part by the fact that if any *S* failed to reach the criterion of learning in 30 trials training was discontinued and he was dropped from the experiment.[7] The instruction to *S* that the egg would be hidden under the "same" block every time may also have contributed to the ease of learning. It should be noted that this instruction may have been taken by *S* either in an absolute or a relative sense, although the question of whether *S* was biased absolutely or relatively by the instructions is not relevant to the results of the experiment. Neither bias could, in itself, contribute to the distance-effect; a relative bias would be expected to facilitate transposition equally at both points, and an absolute bias would preclude transposition at both points.

TABLE I

Experimental Design

Group	N	Training	Test 1	Test 2
I (a)	6	1–2–3	2–3–4	6–7–8
I (b)	6	6–7–8	5–6–7	1–2–3
II (a)	6	1–2–3	6–7–8	2–3–4
II (b)	6	6–7–8	1–2–3	5–6–7

Four of the 24 *S*s (one in each subgroup) were able to point without error to the largest, middle-sized, and smallest blocks in the two sets used in the naming tests. There was, however, not a single *S* who could explain the significance of the middle-sized block for the solution of the problem. With regard to the concept of middleness, therefore, these *S*s may in general be considered 'preverbal.'

On the first trial of T_1, which was a near-test for Group I, 6 of the 12 *S*s in this group chose the intermediate stimulus (Table II). Of the 12 *S*s in Group II, only one chose the intermediate stimulus on the first trial of T_1, which for this group was a far-test. The difference between the two groups in this respect is significant beyond the 4% level by Fisher's exact method.[8] If the two *S*s in each group who made no errors on the naming test are eliminated, the significance of the results is unchanged. Of the 10 remaining *S*s in Group I, 5 transposed on the first trial T_1, while not one of the 10 *S*s in Group II transposed. This difference, by the exact method, is significant beyond the 2-% level. On the first trial of T_2 (far-test for Group I and near-test for Group II), not a single *S* in either group transposed successfully. Thus, the usual distance-effect appeared in the data for T_1, but not in the data for T_2. If within-group comparisons are made instead of between-group comparisons for the first trial of each testing series, similar results are obtained. That is, a significant distance-effect (beyond the 2-% level) appears in Group I (near-test first) but not in Group II (far-test first).

[7] There were five *S*s in this category.

[8] R. A. Fisher, *Statistical Methods for Research Workers*, 1950, 96-97.

The results for all six trials in each test (Table II) show the same pattern as do those for the first trials alone. On T_1, Group I (near-test) made a greater number of intermediate choices than did Group II (far-test), the difference being significant at the 2-% level by Wilcoxon's test for unpaired replicates.[9] On T_2, however, the performance of the two groups did not differ significantly. Within the data for Group I (near-test first), a significant distance-effect appeared (beyond the 2-% level by Wilcoxon's test for paired replicates).[10] The distance-effect did not, however, appear in the data for Group II (far-test first).

The absence of a distance-effect in the data for Group II may be due to the fact that, in one sense, both tests for this group were 'far-tests.' The first test was made with a set of stimuli 'far' from the training set, and the second test with a set of

TABLE II

FREQUENCY OF INTERMEDIATE CHOICE ON THE FIRST TRIAL OF EACH TEST AND MEAN FREQUENCY OF INTERMEDIATE CHOICE ON ALL SIX TRIALS OF EACH TEST

	First trial				All trials			
	Test 1	Test 2	Near −far	Sig.*	Test 1	Test 2	Near −far	Sig.†
Group I	6	0	6	1%	4.50	2.75	1.75	2%
Group II	1	0	1	—	3.08	3.42	0.34	—
Near−far	5	0			1.42	0.67		
Sig.‡	4%	—			2%	—		

* Fisher's exact method.
† Wilcoxon's nonparametric method for paired replicates.
‡ Fisher's exact method for the data of the first trial and Wilcoxon's nonparametric method for unpaired replicates in the case of the combined data for all trials.

stimuli 'far' from the first testing set. For Group I, by contrast, the first test was made with stimuli near the training set, while the second test was made with stimuli far from both the previous sets. In any event, the failure of transposition in the performance of Group II on the near-test provides further evidence for the inability of the *S*s to grasp the concept of intermediate size; even after differentially reinforced experience with two sets of stimuli widely separated on the size-continuum there was no evidence of a generalized intermediate preference.

The relatively high frequency of transposition in the near-test for Group I may be traced in part to the fact that the difference in size among the stimuli was very small. The results of previous experiments suggest that when the differences in size among the stimuli to be discriminated are small, solution of the problem is more likely to require some sort of relational functioning, and the probability of transposition is increased. Furthermore, when the differences within a set of stimulus-objects are small, the differences between successive sets are also small. Transposition in a near-test may, therefore, be due to the fact that *S* is unable to discriminate between the successive sets of stimulus-objects and continues his old mode of response on the testing trials. It may be asserted that, under such condi-

[9] Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures,* American Cyanamid Co., Stamford, Conn., 1946, 1-16.
[10] *Ibid.,* 1-16.

tions, transposition in the intermediate-size problem would be predicted by any theory. To criticize the absolute theory for failing to account for transposition when the difference between sets is small is like saying that a theory is erroneous because it does not explain why *Ss* fail to discriminate differences which are subliminal. The plausibility of this argument derives from the implicit assumption that the behavior of *S* may be determined by certain properties of a set of stimulus-objects as a whole. In fact, however, the absolute theory makes no such assumption. If, after training with the set *1-2-3*, *S* responds to *2* rather than *3* only because the net excitatory value of *2* is greater than that of *3*, he cannot be expected to go to *3* rather than to *2* in the set *2-3-4*. This deduction holds irrespective of the absolute sizes of the stimuli.

Summary and Conclusions

Children between the ages of four and six years were trained to discriminate the intermediate member of a set of three stimulus-objects which differed in size, and then they were tested with sets one and five steps removed from the training set. Despite the inability of the children to verbalize the basis of solution, they transposed significantly to the near set, but there was no transposition to the far set. This distance-effect cannot be traced either to a purely absolute process or to a purely relational ('verbal') one. The solution must have been in part relational (as indicated by transposition to the near set, which, in the intermediate-size problem, cannot be dealt with in absolute terms) and in part absolute (as indicated by the lack of transposition to the far set). This outcome requires a change in the interpretation suggested by previous research on the distance-effect as a function of verbal ability. Available data can no longer be encompassed by a distinction between an abstract, verbally-mediated relational process and a purely absolute one. What is required instead is a distinction between two relational processes—one that is abstract and one that is more closely bound up with the absolute properties of specific situations.

A STUDY OF AUTONOMOUS CHANGE IN THE MEMORY TRACE BY THE METHOD OF RECOGNITION

By JEAN B. CARLSON and CARL P. DUNCAN,
Northwestern University

Of the many studies designed to test the Gestalt theory of autonomous change in the memory trace,[1] one of the best is that of Hebb and Foord.[2] By testing recognition rather than reproduction, and by using different groups of *Ss* at each retention-interval, they avoided common difficulties to which Hanawalt had earlier called attention.[3] The results did not, in their opinion, support the hypothesis of autonomous change. The study by Hebb and Foord has, however, been criticized on two counts. George argued that their longest time-interval (24 hr.) might not have been long enough to permit autonomous changes to take place.[4] He repeated the experiment, extending the retention-intervals to eight days, but again found no support for Gestalt theory. Hilgard pointed out that the figures used by Hebb and Foord, a circle with a gap at the top, and an arrowhead, were not irregular figures, whereas Koffka had argued that the theory should be tested with irregular figures.[5] George substituted ovals for the arrowheads, but ovals, too, are regular figures.

The present study represents a further attempt to test the theory of autonomous change. It follows, in general, the procedure of Hebb and Foord, but there are certain modifications which will be specified later. The study serves also to evaluate the method of recognition which may not be entirely appropriate to the study of autonomous change.

Materials. Two stimulus-figures were used. One was a circle with a 20° gap which was centered at 225° clockwise from the 12-o'clock position. This figure was selected on the assumption that it would create more strongly the impression of an incomplete circle than the one with a gap at the 12-o'clock position which was used by Hebb and Foord. The second figure was an inverted *V*, the right leg of which was 1 1/16 in. long, and the left leg twice that length. It was assumed to meet Koffka's demand for an irregular figure which would create an impression of asymmetry. Each stimulus-figure, drawn in India ink on an 8-in. square of heavy cardboard, was 2 in. in height. The stimulus-cards were larger than those used by Hebb and Foord.

There were 19 recognition-cards for each stimulus-figure. For the circle, the recognition-figures were circles with gaps ranging from 2° to 38° in steps of 2°. For the inverted *V*, the recognition-figures varied in length of the right leg from ½ in. to 1 ⅝ in. in steps of 1/16 in. In both cases the card duplicating the stimulus-figure was the center card of the series. Hebb and Foord used 24 recognition-cards

* Accepted for publication June 11, 1954.

[1] F. Wulf, Tendencies in figural variation, in W. D. Ellis (ed.), *A Source Book of Gestalt Psychology,* 1941, 136-148; Kurt Koffka, *Principles of Gestalt Psychology,* 1935, 493-506.

[2] D. O. Hebb, and E. N. Foord, Errors of visual recognition and the nature of the trace, *J. Exper. Psychol.,* 35, 1945, 335-348.

[3] N. G. Hanawalt, Memory trace for figures in recall and recognition, *Arch. Psychol.,* 31, 1937, (No. 216), 1-82.

[4] F. H. George, Errors of visual recognition, *J. Exper. Psychol.,* 43, 1952, 202-206.

[5] E. R. Hilgard, *Theories of Learning,* 1948, 200.

with the duplicate of the stimulus-card at the eighth position, *i.e.* 7 cards with smaller gaps, 16 with larger gaps.

The recognition-cards also were 8-in. squares of cardboard and were arranged in a continuous series from the smallest gap to the largest for the circles, and from the shortest leg to the longest for the inverted *V*. The cards were loosely bound by two steel rings to form a book.

Subjects. The *S*s were 128 students, both men and women, in psychology classes at Northwestern University. They were divided into 6 groups of 20 to 23 *S*s each. The broken circle was used as the stimulus-figure for Groups 1, 2, and 3, while the inverted *V* was viewed by Groups 4, 5, and 6. Groups 1 and 4 were tested for recognition after 3 min., Groups 2 and 5 after one week, and Groups 3 and 6 after 2 wk. The longest time-interval used by Hebb and Foord was 24 hr., while George extended the longest interval to 8 days.

Procedure. All *S*s were tested individually. The *S* was seated before a table on which *E* placed the stimulus-card, which was covered. The *S* was told that at the signal 'go' the cover would be removed and he was to look at the stimulus-card for 5 sec. Contrary to the procedure of Hebb and Foord, *S* was *not* told that he was to remember the figure exactly or that his memory was to be tested. Immediately after presentation of the stimulus-card, a book was placed before *S* and he was instructed to read until *E* said 'stop.' After 90 sec. *S* was stopped and asked how far he had read. The answer was ostentatiously recorded by *E*. This reading task, not used by Hebb and Foord, was employed in an attempt to disguise in part the nature of the experiment. After the reading period, *S* was either given the recognition-task or told to return 1 or 2 wk. later.

In the test of recognition, the book of cards was placed on the table in front of *S*. The book was opened at different cards for different *S*s. Each *S* was allowed to turn the cards one by one, and was asked to pick out the figure most like the figure he had seen. It is important to note that the series of recognition-cards formed, in effect, an endless belt, with the smallest gap or shortest leg immediately following the largest gap or longest leg. *E* recorded for each *S* both the card at which the series was opened and the card chosen.

Results. In reporting the data, recognition-cards will be identified by reference to their relative position in the series. Cards numbered from −9 to −1 are those with smaller openings or shorter legs than the standard; those numbered from +1 to +9 are those with larger openings on longer legs. The raw score for each *S* is simply one of these numbers, the number of the card he chose.

On the whole, the evidence does not support a theory of autonomous change in the trace, at least insofar as the theory would predict continuous changes in the same direction. Both figures showed reversals in the direction of change, as may be seen in Fig. 1.

Table I gives the mean card chosen for each group and the significance of the difference between each mean and zero, the value assigned the stimulus-card. At the 3-min. interval, there was significant closure of the circle and significant shortening (rather than lengthening) of the leg of the inverted *V*. The table also shows a significant opening of the circle at the 1-wk. interval; no other changes were significant. Although not included in Table I, *t*-values were also computed between each mean and every other for each figure. For the inverted *V*, none of these three

t's was significant. For the circle, the difference between the 3-min. and 1-wk. means was significant at the 1-% level, while that between the 3-min. and 2-wk means was significant at the 5-% level. The 1- and 2-wk. means were not significantly different.

The recognition-method, or at least the variant used here, clearly influenced S's choice. All Ss, except for the 11 who chose the center card of the series, were categorized on the basis of the initial card opened and the card chosen. Of the 58 Ss who began with a negative card, 34 chose a negative; of the 59 who began with a positive card, 36 chose a positive. A test of independence performed on these

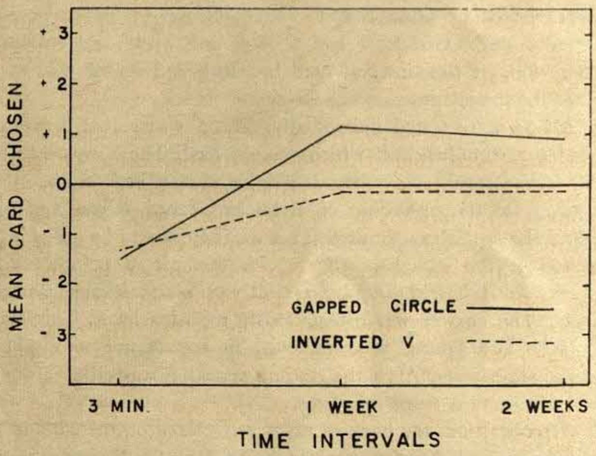


FIG. 1. MEAN CHOICE FOR EACH STIMULUS-FIGURE AND EACH TIME-INTERVAL

frequencies gave a χ^2 of 4.51, significant at the 5-% level. Thus, the place at which the series was opened significantly influenced S's choice.

Essentially the same result can be demonstrated by comparing the mean card (+0.44) chosen by all Ss for whom the series was opened at a positive card with the mean card (—0.83) chosen by all Ss who started at a negative card. The difference between these means is significant at the 5-% level ($t = 2.15$). Since the variances of the groups differed significantly, the method of approximation suggested by Edwards was used to obtain the t-value.[6]

A check also was made to determine if choices were affected by the discontinuity in the endless belt formed by the series—when the series was opened at +9, the next card turned was —9, but if the card first exposed was —9, the next card turned was —8. As is shown in Fig. 2, there was a sharp drop to negative choices by Ss for whom the series was opened at +9, a major deviation from the general trend which may be attributed to the discontinuity in the series.

Discussion. Since the apparent changes in the stimulus-figures over time were not consistent in direction, the data do not support a theory of autonomous changes in the trace. The results cannot, however, be considered crucial for the theory because

[6] A. L. Edwards, *Experimental Design in Psychological Research,* 1950, 167f.

the method used to test for changes in the trace influenced S's choices. The S's response was determined in part by the place at which the series of recognition-cards was opened for him.

TABLE I

MEAN CARD CHOSEN BY EACH GROUP AND THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN EACH MEAN AND ZERO, THE VALUE OF THE STANDARD

Time Interval	Broken circle			Inverted V		
	Mean	t	Signif.	Mean	t	Signif.
3 min.	−1.46	4.42	1%	−1.29	2.08	5%
1 wk.	+1.19	2.09	5%	−0.09	.32	—
2 wk.	+0.95	1.56	—	−0.10	.72	—

With this finding in mind, we may ask whether the experiments of Hebb and Foord, or of George, were conclusive. For the circle with the 20° gap, the mean card chosen was positive for both intervals (5 min. and 24 hr.) in the study by Hebb and Foord, and for three of the four intervals (5 min., 24 hr., and 4 days) in the study by George. The only negative mean was found by George at the eight-day interval. Since a positive mean indicates greater opening of the gap, these authors felt that the Gestalt theory, which would presumably predict closure of the gap, was not supported. In both studies, however, the recognition-figure identical with the stimulus-figure was in the eighth position rather than in the middle of the series of 24 recognition-cards. The series was opened randomly at a card below

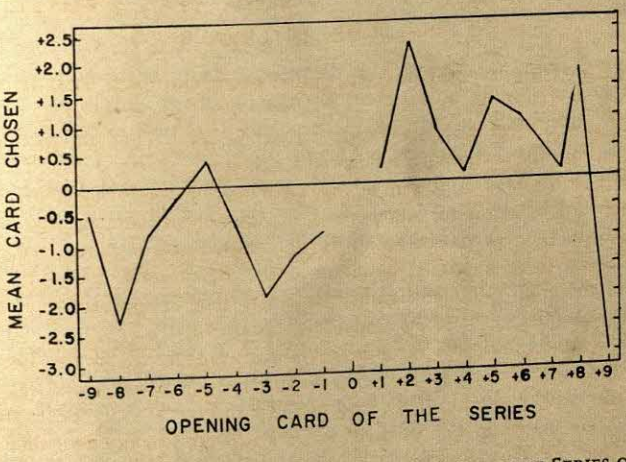


FIG. 2. MEAN CHOICE AS A FUNCTION OF THE PLACE IN THE SERIES OF RECOGNITION-FIGURES AT WHICH THE TEST WAS BEGUN

(The data for both stimulus-figures and all time-intervals are combined.)

eighth position for some Ss, above it for others. Now if, as was true in the present experiment, the Ss tended to choose a card from the portion of the series containing the first card which was exposed in the test, the mean card chosen, based on all Ss, would tend to be positive. Since there were 16 cards above the eighth position and

only 7 below, *S* could choose a card with a larger positive deviation from the eighth card, than the negative deviation of any card below the eighth. Thus, the apparent opening of the gap in the circle reported by Hebb and Foord, and by George, may have been due to the fact that the stimulus-card was off center and the possibility that choices were influenced by the place at which the series was opened. Since the latter error was not controlled in the present experiment, none of the three studies has adequately tested the theory of autonomous change.

Summary. An attempt was made to test the Gestalt theory of autonomous change in the memory trace. The stimulus-figures employed were a broken circle and an inverted *V* with a shortened right leg, and recognition was tested after 3 min., 1 wk., or 2 wk. Separate groups of *S*s were used for each figure and each time-interval.

The results indicated a significant closure of the gap and a shortening of the leg at the 3-min. interval, but these trends reversed at the 1- and 2-wk. intervals. Evidence was obtained which suggests that the method employed in this and in previous experiments is inappropriate.

TIME AS A VARIABLE IN TRANSPOSITION

By HAROLD W. STEVENSON and ERWIN S. WEISS, University of Texas

In his original investigation of transposition Köhler observed that the incidence of transposition may differ, depending upon whether *S* is tested immediately after training, or after a lapse of time following the training trials. For human adult *Ss* he found that "if the test pair is presented (without knowledge) not too soon after the original pair was seen, adults too will, without knowing it, give the same transposition response."[1] Although there have been numerous investigations with animals and children of the effects of other variables, no further exploration of the time-variable has been made. It is the purpose of this experiment, therefore, to assess the generality of Köhler's observations and to investigate the effects of the time-variable in a systematic fashion.

Subjects. The *Ss* were 108 undergraduate students at the University of Texas. All were naïve concerning the type of problem being studied.

Apparatus. The materials were: (a) six 3-in. square blocks each carrying a stimulus-square; (b) a tray for presenting the blocks; (c) a screen to shield *E;* and (d) a collar stud as a goal-object. Mounted in the center of each block was a white cardboard square, the object to be discriminated. The size of the squares on successive blocks increased by 3/16 in., with the smallest square (on Block 1) being 9/16 in. wide. On the tray were four wells in one of which the goal-object could be concealed. The blocks and tray were painted flat black.

Training-procedure. The problem presented to *S* was that of selecting the correct block from a set of four blocks adjacent in the series of sizes. *S* was seated at a table across from *E*. In front of *S* were the tray and the four blocks that were to be used during training. At *S*'s right was the screen. *E* demonstrated how the blocks were to be placed over the holes in the board, and then gave the following instructions: "I am going to hide this collar stud under one of these blocks and I want you to find it as often and as consistently as you can. You will be allowed only one choice per trial."

Following this, *E* moved the tray and blocks behind the screen. After the stud had been placed in one well and the blocks arranged over the proper wells, the tray was again placed in front of *S*. *S* was allowed to pick up one block, then the board was removed and the blocks rearranged. This procedure was repeated on each trial until *S* had made four consecutive correct choices. The arrangement of the blocks on the tray was determined randomly on successive trials but was the same for all *Ss*. For a particular *S*, the goal-object was always under the same block, and this correct block was determined by the experimental conditions.

Test-procedure. A changed set of blocks was used on the test-trials. The set differed from the training-set in that one or two blocks at one end of the training-series were removed, and an equal number of new blocks from the opposite end of

* Accepted for publication July 26, 1954.

[1] Wolfgang Köhler, Simple structural functions of the chimpanzee and the chicken, in W. D. Ellis (ed.), *Sourcebook of Gestalt Psychology*, 1938, 225.

the size-series was added. The correct square was defined as the one with the same ordinal position in the test-series as the correct one in the training-series. For instance, if the goal-object was under the second largest square in the training-set, it was under the second largest square in the test-set. The *Ss* were trained to choose the correct square in the test series and continued until they had made four consecutive correct responses.

Experimental groups. There were three groups in the principal experiment. Group I was tested immediately after training; Group II was tested after 10 min.; and Group III was tested after 24 hr. Two sets of blocks were used: Set A consisted of Blocks 1, 2, 3, and 4; and Set B consisted of Blocks 3, 4, 5, and 6. Half of the *Ss* in each group were trained on Set A and tested on Set B, and half were trained on Set B and tested on Set A. On the training-trials Block 3 was correct for half of the *Ss* and Block 4 was correct for half. This made it possible to shift the blocks two steps and still have the originally correct block present on the test-trials.

An additional group, Group IV, was run with immediate testing, but with the test-set of blocks shifted only one step from the training-set. In this group all *Ss* were trained on Blocks 2, 3, 4, and 5. A shift was made one step up (to Blocks 3, 4, 5, and 6) or one step down (to Blocks 1, 2, 3, and 4), depending upon whether Blocks 2 or 3 or Blocks 4 or 5 were correct during training. Again, the correct block during training was present throughout the test-trials.

In the immediate testing of Groups I and IV no comment was made by *E* to introduce the test-trials and the trials were begun as soon as the training-criterion was reached. In Groups III and IV, *S* was asked to return after the appropriate time-lapse to complete the experiment. Upon returning he was told: "We'll continue with what we were doing."

Verbal reports. All pertinent reports during the experiment were recorded. If *S* did not say anything about his performance, he was asked at the conclusion of the test-trials about his method for finding the stud. If he failed to mention whether he noticed the change in blocks, he was asked if the blocks in the test series which were still before him were the original blocks. If *S* recognized the change, he was questioned concerning the basis of his first test-choice.

Results: (1) Training-trials. The number of trials required to reach the training criterion (excluding the four criterion-trials) ranged from 0 to 113. In all groups the distribution of trials was highly skewed. Comparisons of the means (see Table I) by use of the *F*-test reveal no significant differences.[2]

(2) Test-trials. The discrimination was learned so rapidly on the test-trials that after the initial trial, differences between the groups quickly disappeared. Consequently, the measure of transposition adopted is the proportion of relational (correct) choices on the first test-trial. According to this measure, shown in the last column of Table I, all groups except Group III transposed, for the proportion of relational choices was significantly above chance ($CR > 1.96$; $p < .05$).

The incidence of transposition differed, depending upon the testing conditions. With blocks two steps removed, there was significantly more transposition following a delay of 10 min. than following immediate testing ($\chi^2 = 4.80$; $p < .05$), and

[2] Leon Festinger, A statistical test for the means of samples from skew populations, *Psychometrika*, 8, 1943, 205-210.

significantly less transposition after 24 hr. than after 10 min. ($\chi^2 = 6.40$; $p < .05$). There was significantly less transposition when the blocks were two steps removed than when they were only one step removed ($\chi^2 = 4.80$; $p < .05$).

(3) Verbal reports. A significant negative relationship was found between the incidence of transposition and the *S*s' verbal reports about noticing the differences between the sets of training- and test-stimuli. Of the 50 *S*s in all groups who stated that the sets were different, 26 transposed on the first test-trial, while of the 58 *S*s

TABLE I

PERFORMANCE ON TRAINING- AND TEST-TRIALS

Group	N	Number of training-trials M	Number of training-trials SD	Number of test-trials M	% correct first test-trial
I	30	17.5	19.2	0.7	53
II	30	19.1	15.5	0.4	80
III	18	18.3	26.8	3.1	44
IV	30	18.4	20.0	1.6	80

who stated that the sets were the same, 46 transposed. The difference is significant at greater than the 0.01-level of confidence ($\chi^2 = 9.01$).

A breakdown of the data by groups (Table II) reveals that significantly more *S*s noticed the difference when the testing was done immediately than when it was delayed for 10 min. ($\chi^2 = 31.09$; $p < .001$). The delay of 24 hr. did not produce significantly different effects from the delay of 10 min. ($\chi^2 = .007$). When only one new block was introduced on the test-trials significantly fewer *S*s noticed the change than when two new blocks were introduced ($\chi^2 = 31.09$; $p < .001$).

A statistical analysis of the combined interaction between experimental condition, noticing the change, and the incidence of transposition is prohibited by the small

TABLE II

PERFORMANCE ON THE FIRST TEST-TRIAL AS A FUNCTION OF REPORTED NOTICE OF A CHANGE IN SIZE BETWEEN TRAINING- AND TEST-STIMULI

	Group I	Group II	Group III	Group IV
Size-change noticed; relational test-response	15	4	2	5
Size-change noticed; non-relational test-response	14	4	3	3
Size-change not noticed; relational test-response	1	20	6	19
Size-change not noticed; non-relational test-response	0	2	7	3

number of *S*s in some of the cells of Table II, but an examination of the table suggests the presence of such an interaction. The *S*s who reported that the blocks were different made approximately an equal number of relational and non-relational responses, while the *S*s (except in Group III) who did not notice the change in the blocks consistently transposed at a high level.

The divergent results for the *S*s in Group III may be related to the fact that of the nine *S*s trained on the end block, seven transposed; while of the nine *S*s trained on one of the middle blocks, only one transposed. This difference is significant ($\chi^2 = 5.62$; $.01 < p < .02$), while the difference for the combined data from the other groups is not ($\chi^2 = .55$). Thus the *S*s may have failed to notice the size-change

and still not have transposed at a high level because of the difficulty of retaining for 24 hr. the more complicated relationships among the blocks when either of the middle ones was correct.

Discussion. These results are of interest theoretically in that a relational theory such as that of Köhler can explain, in the fashion outlined earlier, the improved transposition when a short delay is introduced between the training- and test-trials.[3] The discrimination theory of Spence, on the other hand, has not considered the effects of time on transposition but has found strong support in its ability to predict the decrease in transposition with increasing remoteness of test-stimuli.[4] Neither Spence nor Köhler has, however, discussed explicitly the possibility that transposition may be due in part simply to the inability of *S*s to discriminate between the training- and test-sets. Such a factor has been mentioned in previous discussions of transposition, and the results of this experiment indicate clearly that even though *S*s are capable of discriminating among the members of a set of stimulus-objects presented simultaneously, under certain conditions they may fail to discriminate between consecutive sets.[5] In such cases transposition is practically assured, because the failure to discriminate between the sets of stimulus-objects results in a high probability that *S* will respond to the test-objects in the same relative fashion as he responds to the training-objects.

Summary. College students were trained to choose one of four stimulus-objects differing in size, and after different intervals of time the *S*s were tested on stimulus-objects either one or two steps removed from the training-set, *i.e.* one or two new objects were introduced on the test-trials. Transposition was most frequent when *S*s were tested immediately with objects only one step removed and when they were tested after 10 min. with objects two steps removed. It was less frequent when the *S*s were tested with objects two steps removed either immediately or after 24 hr. The *S*s' verbal reports indicated that some *S*s failed to notice that the training- and test-sets of objects were different. The difference was noticed least frequently when objects one step removed were presented immediately and when objects two steps removed were presented after 10 min. The relationship between noticing the change and transposition was highly significant statistically, with the *S*s who failed to notice the change showing more frequent transposition.

---

[3] Köhler, *loc. cit.*

[4] K. W. Spence, The differential response in animals to stimuli varying within a single dimension, *Psychol. Rev.*, 44, 1937, 430-444.

[5] R. H. Gundlach and G. B. Herington, The problem of relative and absolute transfer of discrimination, *J. Comp. Psychol.*, 16, 1933, 199-206; J. A. McGeogh, *The Psychology of Human Learning*, 1948, 433-435.

THE RELIABILITY OF OLFACTORY THRESHOLDS OBTAINED BY SNIFFING

By F. NOWELL JONES, University of California, Los Angeles

Although it is possible to present odorous stimuli in a controlled manner by the use of special apparatus, requirements of portability, speed, and simplicity make a straightforward sniffing technique appealing.[1] The first step in appraising such a method is the determination of its reliability, especially since many of the criticisms leveled at sniffing methods have emphasized the variability from sniff to sniff. If sniffing should prove reasonably reliable, the further effort to refine it would be in order.

Materials. Three series of odors were prepared in ordinary wide-mouthed 200 ml. bottles. The series were obtained by successive dilution with mineral oil as follows: (1) n-butanol, starting with a 0.021 molar solution, and diluting 1 part of preceding level with 1 part mineral oil, for a total of 23 steps; (2) safrol, starting with 0.014 molar solution, and diluting 3 parts of preceding level with 2 parts of mineral oil, for a total of 15 steps; and (3) n-butyric acid, starting with a 0.0038 molar solution and diluting 2 parts of the preceding level with 1 part of mineral oil for a total of 30 steps. These dilutions were kept covered with caps of aluminum foil, and were stored in the experimental room which showed an extreme deviation of 2°F, during the course of the experiment. In addition to this temperature control it may be noted that the room was kept from becoming excessively odorous by the constant operation of a charcoal-air purifier.[2]

Experimental design. In all, 24 *O*s were tested. Most of them were advanced undergraduate psychology majors, while a few were graduate students. The number of *O*s arose from the design of the experiment. To eliminate as far as possible the effect of one odor on another, all six possible permutations of the three odors were drawn up. These six groups of three were then arranged in serial order to give a balanced random series, except that it was considered undesirable to have a substance follow itself immediately and so the resulting order was not perfectly balanced. To eliminate order as far as possible, each successive *O* entered the series at the beginning of a successive group, there being, therefore, six different entry points in the series with four *O*s entering at each one. Since every *O* was tested for the total series, there were six threshold determinations for each of the three substances.

Procedure. The actual method of presenting the stimuli was kept as simple as possible. *O* was seated and given a sample sniff of each odor to ensure easy identification. He was then given a chance to sniff a bottle containing pure mineral oil, which served throughout as a 'zero' reference point in case of doubt. For sniffing, the bottle was held by *E* under *O*'s nose, and *O* was instructed to sniff the contents. In getting the thresholds the ascending method of limits, as is traditional in chemo-

* Accepted for publication June 3, 1954. This study was made possible by a grant from the Research Committee, University of California, Los Angeles.

[1] F. N. Jones, An olfactormeter permitting stimulus specification in molar terms, this JOURNAL, 67, 1954, 147-151.

[2] This purifier was kindly loaned by the Pure Air Division of the American Solvent Recovery Corporation, Columbus, Ohio.

receptor work, was used throughout. Recognition was required. All bottles were kept covered when not actually being sniffed. The total series of 24 thresholds usually occupied 20 to 25 min., a time sufficient to prevent any particular drift in thresholds through adaptation, as will be seen when the results are discussed.

Results. For purposes of obtaining reliability estimates the data were retained in their raw form, that is, by the number of the bottle to which the positive response was given. Since the bottles were numbered sequentially from highest concentration to lowest, this in effect resulted in a ratio transform of each according to the method of dilution. The thresholds for the separate substances were assembled to permit an

TABLE I

THE RESULTS OF ANALYSIS OF VARIANCE, AND MEDIAN THRESHOLDS AND Qs

Substance	$r_{1.1}$	Inter-*O* *F*	Inter-trial *F*	Med. threshold molar ratio	
n-butanol	.82	28.54 $P<.001$	1.68 $P>.05$	2.86×10^{-9}	1.03×10^{-6}
safrol	.77	21.5 $P<.001$	.14 $P>.05$	6.35×10^{-9}	1.14×10^{-8}
n-butyric acid	.80	258.49 $P<.001$	10.77 $P<.001$	1.12×10^{-9}	4.44×10^{-9}

analysis of variance taking account of both *O* and order. Estimates of $r_{1\cdot1}$ for each substance were obtained by means of intra-class correlation, according to the formula for an estimate based on single classification.[3] In addition, *F*-ratios for inter-*O* and inter-trials variance were calculated. These results are summarized in Table I. As a matter of interest, although not directly germane to the present purpose, the median threshold in molar terms and *Q* for each substance are also given. These thresholds refer to the air in the dilution bottles, not the mineral oil solutions.

Discussion. The reliabilities are rather high. In each case, about 80% of the variance is common to the different trials for the same substance. This result is also reflected in the inter-*O* *F*-ratios which are all highly significant, and by the fact that only one inter-trial *F*-ratio was significant, and that at a lower level. It is apparent that taking so many thresholds in succession has had but little effect in comparison with the differences among our *O*s.

Since reasonable reliability has been established, the task remains of determining the bases of the consistent individual differences found. Partly this can be accomplished by comparison with the results of another method, better controlled. Such work is in progress. It will also be desirable to try to improve the sniffing method itself to yield better control. In the meantime, we may conclude that whatever other faults sniffing may have as a way of getting thresholds, variability from trial to trial due to random error is not one of them.

Summary. To determine the reliability of olfactory thresholds obtained by sniffing, thresholds were obtained by this technique from 24 *O*s, for 3 substances. Six thresholds were obtained for each substance by an ascending method of limits. The threshold reliabilities calculated by analysis of variance were: n-butanol, 0.82; safrol, 0.77; n-butyric acid, 0.80. From these results the conclusion seems warranted that sniffing is an adequate technique for threshold studies.

[3] Quinn McNemar, *Psychological Statistics*, 1949, 280.

THE EFFECTS OF DIFFERENTLY STRUCTURED VISUAL FIELDS ON THE PERCEPTION OF VERTICALITY

By MELVIN WEINER, Worcester State Hospital

The present experiment was designed to study the effects of three visual frameworks of varying complexity on the perception of verticality. The simplest framework was a luminous rod, the second in order of complexity was a luminous square, and the most articulated was a luminous cube. The hypothesis tested was that the influence of the visual framework increases with degree of articulation.

Method. The rod and square were identical with those used by Witkin and Asch.[1] The rod was 39 in. long and 1 in. wide. Each side of the square was 40 in. long and 1 in. wide. The third framework, a cubical arrangement, was developed for this experiment to provide a three-dimensional visual field, and gave the appearance of the outline of a cube with the front part missing. The back section was identical with the square, and its sides were strips of wood, 47 in. long and 1 in. wide, which projected toward *S* from each corner of the square. Each visual field was a separate apparatus pivoted at eye-level on a horizontal shaft 8 ft. in front of *S*, and could be rotated and viewed independently. All three structures were coated with white luminous paint, and during the experiment were the only items visible in the otherwise completely darkened room. A protractor and stationary pointer permitted direct readings (in degrees) of the tilt of the rod, square, or cube. A well-padded tilting chair with a high back, side-supports, and foot-rests, and an adjustable sponge-rubber head-rest, allowed *S* to be tilted in various positions.

Each *S* was brought into the experimental room with his eyes closed and masked aviator's goggles over his eyes. He was seated in the upright chair and asked to lift his goggles and open his eyes on each trial. His task was to direct *E* how the visual framework should be adjusted from an initially tilted position to look upright, *i.e.* "straight up and down like the walls of the building or the flagpole outside." Each of the adjustments was preceded by a 3-min. period of observation. Between trials *S* was instructed to keep his eyes closed and lower the mask over them while the protractor was read. Two body-positions (erect and 28° to the left) and two framework-conditions (framework tilted 28° to the left or right) were employed. The order of trials was (*1*) body erect with the framework alternately tilted *L, R, L, R,* and (*2*) body tilted to the left with the framework alternately tilted *L, R, L, R, L, R, L, R.*

Three groups of *Ss* were used for the three different field conditions; 14 *Ss* were tested with the rod, 15 *Ss* with the square, and 12 *Ss* with the cube. All *Ss* were Brooklyn College students who had volunteered for the experiment.

Results. The amounts by which each *S*'s adjustments of the rod, square, or cube

* Accepted for publication July 29, 1954. From the psychological laboratories of Brooklyn College. The author is indebted to Dr. H. A. Witkin and to Dr. Alfred L. Baldwin for helpful advice.

[1] H. A. Witkin and S. E. Asch, Studies in space orientation: IV. Further experiments on perception of the upright with displaced visual fields, *J. Exper. Psychol.*, 38, 1948, 762-782.

deviated from the true upright were averaged without regard to sign, and then means were obtained for all *S*s in each group. Table I presents the mean amounts by which the adjustments for the different frameworks deviated from the true upright.

With *S* erect, greater complexity of structure led to progressively greater errors of judgment. The *t*-test showed the difference between each pair of means to be significant beyond the 1-% level. The upright position of the body makes for the most accurate determination of the gravitational upright, since all the task requires is lining up the field with the upright body. Accordingly, the mean error of 14.7° with the cube is especially noteworthy; the cube was perceived as upright at a point closer to the initial tilt of the visual field than to the bodily vertical.

With the body tilted to the left and the field initially tilted to the same side, the data again show increasing deviation from the gravitational vertical with increasing complexity of structure. The difference between the means for the rod and the square was significant beyond the 5-% level, the difference between the means for the square and the cube significant beyond the 1-% level.

TABLE I

Mean Error of Adjustments (Degrees)

Condition	Rod	Square	Cube
S erect, field *L*+*R*	3.7	7.7	14.7
S left, field *L*	5.6	10.3	19.9
S left, field *R*	18.3	16.1	21.9

The direction of errors in judging the upright with less complicated structures is not random in character, but is systematically related to the position of the body. With the body tilted 28°, an upright rod is perceived as displaced toward the body (*E*-effect).[2] There is, therefore, a strong tendency to move the rod opposite the tilted body to make it look upright. Larger errors are consequently produced when the rod is initially tilted opposite the body than when it is tilted to the same side.

A progressive decrease in the *E*-effect with increase in the complexity of the field was noted in the present experiment. The rod was displaced 12.7° more when it was initially tilted opposite the body than when it was initially tilted with the body, thus indicating a strong *E*-effect. The square was displaced 5.8° more when it was opposite the body; the cube 2.0 more. Deviations of the cube from the true upright were about equally great whether the body and the visual field were tilted in the same or in opposite directions.

The greater freedom from the *E*-effect with more complex visual structures may account for the apparent reversal of the trend in mean errors with the rod and the square, when body and framework were oppositely tilted. Under these conditions the means for square and cube differed significantly beyond the 5-% level, but the means for rod and square did not differ significantly. Actually, however, there was a smaller mean error with the square, which may reflect greater influence of the more complex structure. When the rod is initially tilted opposite to the tilted body, both the *E*-effect and the righting-effect from continued observation work in the same

[2] Witkin and Asch, Studies in space orientation. III. Perception of the upright in the absence of a visual field, *J. Exper. Psychol.*, 38, 1948, 603-614.

direction and large deviations are, in consequence, produced with the rod. With the more complex square, the field-effects are more autonomous. Acceptance of the tilted frame is not supplemented by the *E*-phenomenon, and slightly smaller errors result. Thus, smaller errors with the square may reflect greater freedom from the *E*-effects and greater influence of the framework.

There were important individual differences depending on the visual field. While some *S*s tended to judge the vertical predominantly on the basis of visual cues, others adjusted the fields to positions intermediate between the postural and visual determinants, and still others responded primarily on a postural basis. In addition to individual differences in magnitude of errors, there were also significant differences in qualitative aspects of performance, such as confusion, indecision, dizziness, headache, and disorientation.

Conclusion. In the judgment of verticality, the influence of the visual framework increases with its complexity.

RELATIVE SIZE VS. FAMILIAR SIZE IN THE PERCEPTION OF REPRESENTED DEPTH

JULIAN E. HOCHBERG and EDWARD MCALISTER, Cornell University

The monocular cue of distance—familiar size—has recently received considerable attention due largely to concerted attempts to 'explain' the perception of depth—and in general, all perception of the world around us—in terms of learning and past experience.[1] This empiristic enterprise, which is by no means new to the history of perception, is certainly one legitimate approach to the problem but the evidence in favor of it is still far from conclusive. Regardless of whether a particular 'cue' is ultimately a product of past experience, we can learn much more about its operation and its limitations, if we attempt to relate its present effects to its stimulus-characteristics, than if we dismiss the measurable characteristics of the stimulus and simply ascribe its power to the life-history of the observer.[2] Since familiar size is the only cue of depth that requires, by its very definition, reference to past experience, it seems important to determine its applicability and limitations. This is particularly true since, as has been suggested elsewhere, another monocular distance cue has frequently been confounded with familiar size; namely, that of 'relative size,' a cue which need not invoke past experience.[3]

Familiar size requires a complex array of intervening variables for its explanation. Probably the simplest version would be that a given retinal image will, by its shape and size, evoke a trace or a memory of a specific object at a specific distance, and this trace then determines the distance at which the object is perceived. One retinal image is all the stimulus that is required for the operation of this cue.

Relative size, on the other hand, requires two similar or identical shapes of different size; when that occurs, the larger tends to appear nearer than the smaller. Here a pair of figures is required, and the pair must be similar or identical. This explanation does not refer to past experience; it simply states that images of the same shape but of different size are stimuli for a depth relationship.

Very many of the experiments purporting to demonstrate the operation of familiar size in 'immediate' perception (as opposed to judgment, reflection, or whatever) are open to reinterpretation in terms of relative size.[4] A playing card casting a large retinal image is seen as being nearer than a playing card casting a small one, a situation which has been repeatedly interpreted as evidence for familiar size and, therefore, for past experience as a determinant of immediate perception of depth.

---

* Accepted for publication August 21, 1954.

[1] A. H. Hastorf, The influence of suggestion on the relationship between stimulus-size and perceived distance, *J. Psychol.*, 29, 1950, 195-217; C. B. Hochberg and J. E. Hochberg, Familiar size and the perception of depth, *J. Psychol.*, 34, 1952, 107-114; Hochberg and Hochberg, Familiar size and subception in perceived depth, *J. Psychol.*, 36, 1953, 341-345.

[2] J. E. Hochberg, Psychophysics and stereotypy in social perception, *Emerging Problems in Social Psychology* (in press).

[3] Hochberg and Hochberg, *op. cit.*, *J. Psychol.*, 36, 1953, 343.

[4] W. H. Ittelson and F. P. Kilpatrick, Perception, *Sci. Amer.*, 185, 1951, 50-56. Hochberg and Hochberg, *op. cit.*, *J. Psychol.*, 36, 1953, 342.

Before we accept this as evidence for such a position, it should be noted, however, that such results may be perfectly well explained in terms of the relative size. Because of the considerable difference between the two cues, it seems necessary to determine further whether such a cue actually exists as separate from familiar size.

In the present experiments, we ask whether relative size operates, as a separable cue in the *representation* of depth, in simple two-dimensional line drawings.

Experiment 1. In our first experiment, 4 stimulus-cards, 40 in. wide and 30 in. high, were presented in balanced order to 24 *S*s. Each card bore two figures, one small one and one large one, 8 in. apart. These are the figures that appeared on the 4 cards: Card 1 had a large circle and a small circle; Card 2 had a large square and a small square; Card 3 had a large circle and a small square; and Card 4, a large square and a small circle. In each case, the large figure was 16 in. across, while the smaller figure was 4 in. across. Each stimulus-card was presented for 100 sec.; the *S*s were shown the possible two-dimensional and three-dimensional phases of the line drawings, and then instructed to indicate whether, with a passive attitude, the two figures appeared in the same or different planes. Thirty-three signals (sounds) randomly spaced were given during the presentation and the *S*s were instructed to

TABLE I

DEPTH RESPONSES TO PAIRS OF GEOMETRIC STIMULUS-FIGURES

Experiment		Stimulus-cards			
		1	2	3	4
1	% 3-dimensional responses	56.5	58.5	38.5	40.5
	Means, 1+2, 3+4	57.5		39.5	
	Significance of difference	P<.01, N=1980			
2	% large figure nearer	54.2	64.4	51.9	40.4
	Significance of difference from 50%	P<.01	<.01	—	<.01
	No.=1710				

indicate the mode of appearance of the figure which *immediately preceded the signal.* Thus, the relative perceptual response-probabilities were indicated by the random time-intervals taken, and proportions of two- and three-dimensional responses were obtained for the stimulus-cards.[5]

In terms of a cue of relative size, Cards 1 and 2 should yield more three-dimensional responses than Cards 3 and 4 (in which the stimuli are pairs of figures of different shape). As Table I shows, they clearly did. Even if the concept of familiar size could apply to geometrical line drawings, we cannot in any case readily apply it alone to explain these results, because if it were at all applicable here, it should be just as applicable to Cards 3 and 4, as to Cards 1 and 2, since it should operate with respect to each figure in the pair irrespective of their relationship. We must now inquire whether the direction of the three-dimensional responses are in accordance with what would be predicted in terms of relative size.

Experiment 2. In Experiment 2, the procedures and stimulus-cards of the first experiment were repeated, with new groups of *S*s (83 in number). These were

[5] J. E. Hochberg and E. McAlister, A quantitative approach to figural "goodness," *J. Exper. Psychol.,* 46, 1953, 361-364.

given the additional instructions to indicate within the three-dimensional phases, whether the left or the right figure appeared to be nearer.

In terms of the cue of relative size, the larger figure should appear nearer than the small one in Cards 1 and 2. They did. If this were due to the operation of familiar size, we would expect at least similar results to hold with respect to Cards 3 and 4, but again, as Table I shows, they did not—in fact, for Card 4, the relationship is significantly reversed.

For the two shapes employed, we can say that figure pairs of the same shape, but different size, are cues for the perception of a difference in depth, with the larger of the two appearing to be nearer.

In short, relative size seems to be an at-least formally separable monocular cue of depth, which, aside from its own predictive applications, must be carefully separated from any situation in which experimenters are attempting to demonstrate the operation of past experience in perception.

APPARATUS

AN ELECTRONIC APPARATUS FOR STUDYING TRACKING PERFORMANCE

By A. A. GERALL, P. B. SAMPSON, and S. D. S. SPRAGG,
University of Rochester

This note describes the design and operation of a two-hand tracking apparatus which is at present being used in a study of simple tracking but is adaptable to continuous, compensatory tracking. The device consists of the following units: (1) subjects' controls; (2) display; (3) programmer;

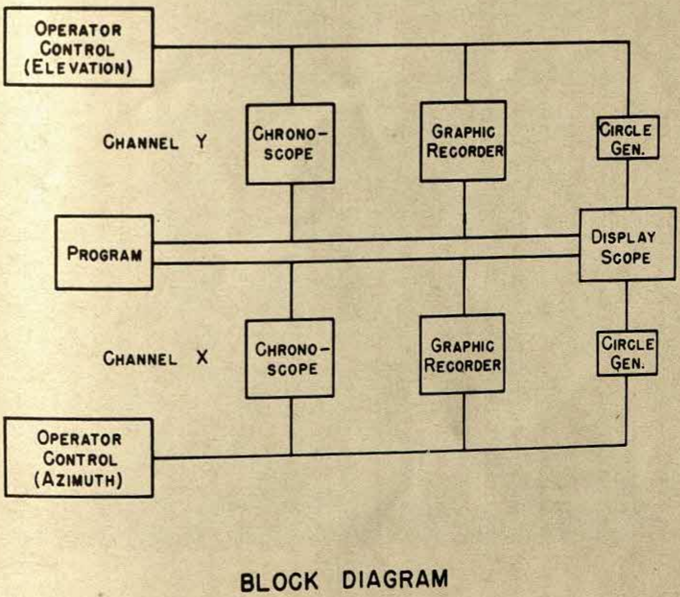


FIG. 1. BLOCK DIAGRAM OF THE CONTROL, DISPLAY, AND RECORDING UNITS

and (4) recorders. Fig. 1 presents a block diagram of the functional connection of these units and Fig. 2 illustrates in schematic form the electronic details of the circuits. The operation of the apparatus is explained and the essential details of its design and construction are described.

* The design and construction of this equipment was carried out under Contract No. DA-49-007-MD 326, Surgeon Generals Office. This is part of Report No. 144 under AMRL Project 6-95-20-001. The authors would like to thank Dr. Charles Dawson and Mr. Gordon Gay for their assistance in construction of this apparatus.

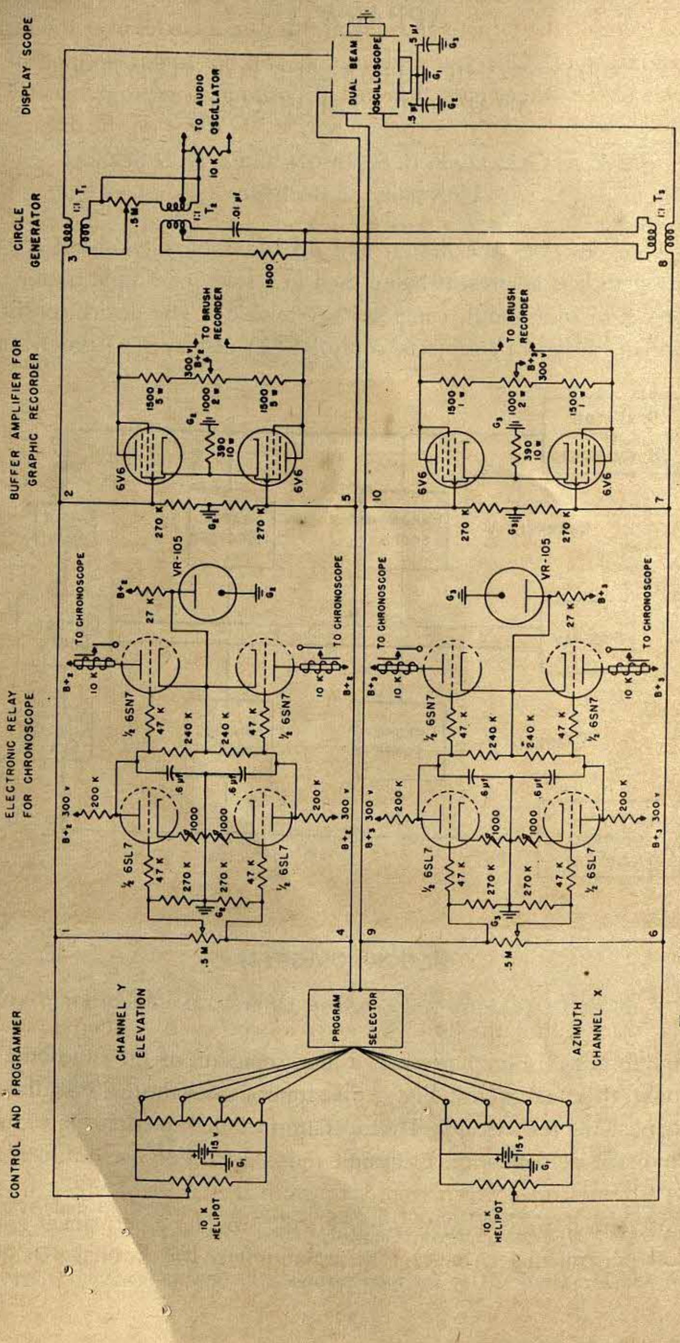


FIG. 2. SCHEMATIC DIAGRAM FOR THE CONTROL, DISPLAY AND RECORDING CIRCUITS

Two bar-cranks, as shown in Fig. 3, are available to *S* for controlling the movement of the target-follower, in this case a circle, over the face of a cathode ray tube. One crank determines the horizontal position of the follower and the other crank the vertical position. The programmer is designed to present the target, a spot, at various locations on the display unit

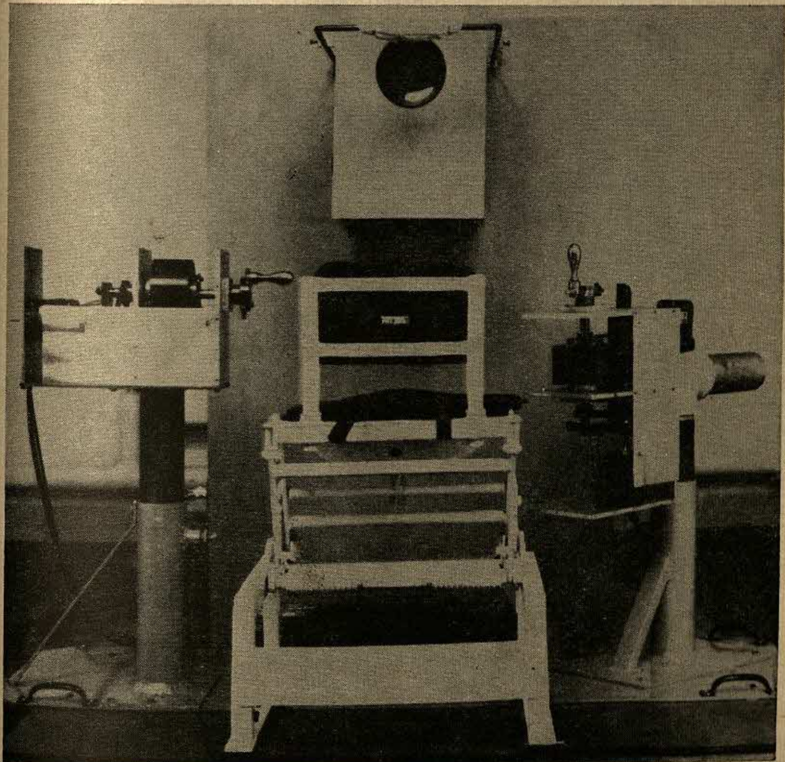


FIG. 3. OPERATOR'S SIDE OF THE APPARATUS

—a dual-beam cathode ray tube (*CRT*). When the target appears, *S* rotates the cranks until the follower encircles the spot. A short time after he accomplishes this simple task, which is sometimes referred to as 'target acquisition,' the follower and target disappear from the *CRT* until a new displacement is presented. Two circuits are included to obtain information concerning *S*'s performance. Electronic relays are used to activate Standard Electric Chronoscopes which record 'time-off' the target in either the horizontal or vertical directions. A second circuit provides a permanent graphic record of performance. A voltage corresponding to the difference in spatial

location between the follower and target is amplified and fed into a Dual Channel Brush Oscillograph. The direction and magnitude of *S*'s errors are continuously registered on chart paper moving at a constant speed. *E*'s only task during an experimental period is to record 'time-off' the target from the chronoscopes. In the following paragraphs a more detailed discussion of the components of the apparatus will be presented.

(1) Control unit. The control unit consists of two adjustable bar-cranks which can be located in several positions and planes. Rotation of a crank governs the movement of the follower in either the horizontal or vertical direction on the *CRT*. The cranks are coupled to the shafts of electromagnetic clutches. By varying the current passing through the coils of the clutches, different amounts of friction—predominantly coulomb friction—are obtained. Approximately 2 to 90 in.-lb. of torque can be secured with these units. The movement of the shaft is transferred by means of a pulley system to a helipot and a protective stopping unit. This unit locks at the end of 12 turns in either direction from the center of the helipot. If the crank is rotated more than 12 turns from the center, the locking unit causes the pulley from the crank shaft to the helipot to slip, thereby permitting *S* to continue turning the crank without extending the sliding arm of the helipot beyond its end-point. A frontal view of the control unit described above is shown in Fig. 4.

(2) The display. The display is presented on the face of the cathode ray tube of a DuMont Type 322 Dual-Beam Oscillograph. One beam of the tube is focused to a spot and is employed as the target. *S*'s task is to place the follower over the target by the appropriate rotation of the cranks. The circuits used to achieve this display are shown in Figs. 2 and 6.

The location of the target or follower on the scope is determined by the D.C.-potential available to the oscillograph. The vertical displacement of both the follower and the target is determined by a parallel resistance network. An identical network controls the horizontal location of both the target and the follower. This network consists of a 10,000 Ω potentiometer in parallel with nine resistors in series equalling 10,000 Ω, and a 15-v. center-tapped D.C.-supply. Hence, by moving the arm of the helipot or by selecting a particular tap along the series of fixed resistors different potential differences in respect to ground can be obtained. These D.C.-differences are coupled directly to the oscillograph. If a resistor above ground in the 'y' channel of the oscillograph is selected, the target appears above the center of the *CRT* since the center of the D.C.-supply is connected to

the ground of the scope. A potential difference would also exist between the lead from the helipot and the lead from the fixed resistor and this would appear on the *CRT* as a discrepancy in the vertical position between the target and follower. When the sliding arm of the helipot is moved to a position at which the resistance from the arm to ground equals

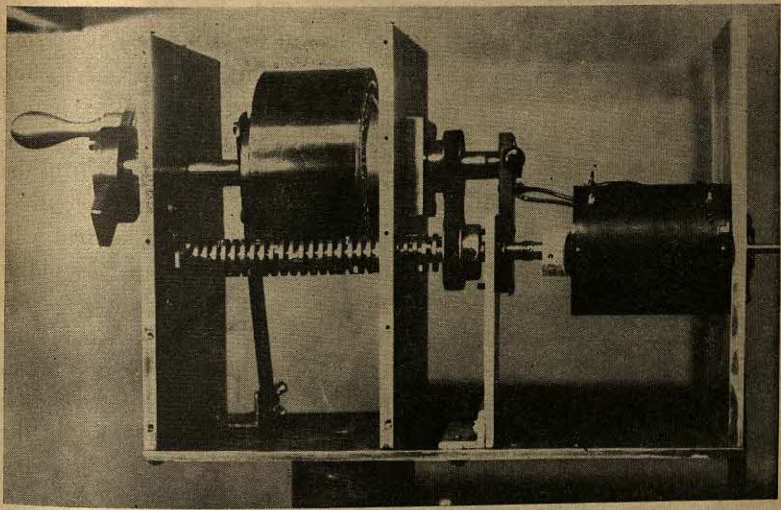


FIG. 4. CONTROL UNIT
The main parts are the helipot, locking unit, electromagnetic clutch, and bar-crank unit.

the companion resistance of the parallel network, the potential difference between these input leads to the oscillograph becomes zero. On the face of the *CRT* the target and follower would be in the same position in the vertical direction. The same description, of course, would apply to the network used to control the displacements in the horizontal direction.

A circle generator is placed along the leads from the helipots to the oscillograph. The generator consists of an audio oscillator which produces a 3500 cycles/second signal and a R.C. phase-splitting network. The outputs of the R.C.-network are two identical voltages except that one is 90° out of phase with the other. By means of transformer coupling one of these voltages is transferred to the lead from *S*'s circuit going to the horizontal deflection plate of the oscillograph and the other voltage to the vertical deflection plate. Hence the D.C.-signal from *S*'s circuit is modulated by two alternating voltages which are 90° out of phase with each other. This provides a Lissajous figure which is a circle for this phase dif-

ference. Also, the modulating voltage has no influence upon the position of the circle on the *CRT* which is solely determined by the D.C.-signal from the helipot in *S*'s circuit.

(3) Programmer. The programmer determines the location or displacement of the target on the face of the *CRT*. It also controls the inter-trial

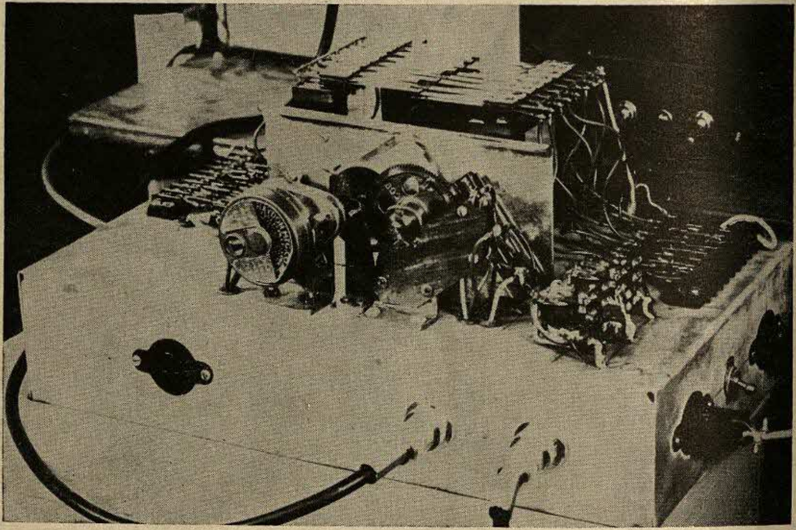


FIG. 5. PROGRAMMER

interval, the blanking of the scope and the operation of the Brush Oscillograph and chronoscope motors. Fig. 5 shows the physical construction of the unit and Fig. 6 presents a schematic representation of the electrical components. A point along the series of fixed resistors is selected by means of cam-operated lever-switches. A lever-switch in the normally open position is connected to each of eight positions on this resistance network. A given resistor is selected when a cam closes its lever-switch and connects it to the circuit of the oscillograph. The rotation of the cams is controlled by a constant-speed motor. Hence eight different displacements can be presented successively at fixed intervals. A grooved drum is mounted on the same shaft to which the cams are connected. Microswitches balanced on this drum operate the blanking relays and the A.C. to the Brush and chronoscope motors. When a lever switch is disengaged from a cam, a microswitch closes and deënergizes two relays. Then the contacts of these relays are released and all inputs to the oscillograph are opened leaving only

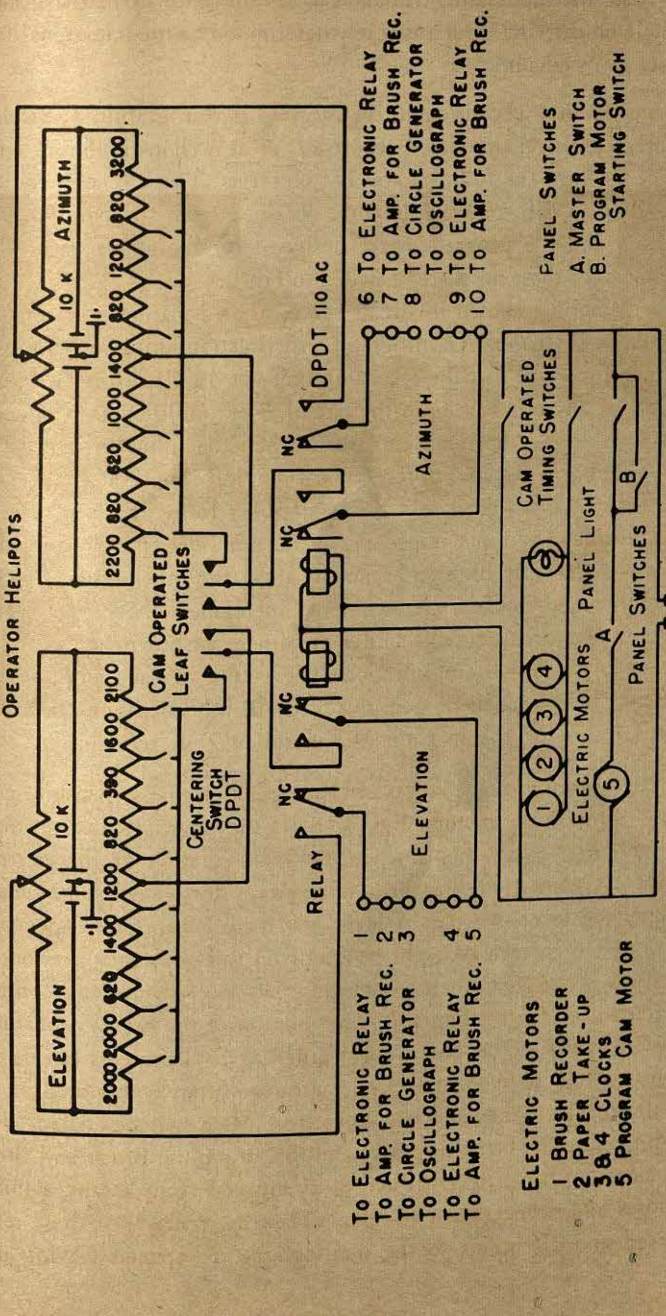


Fig. 6. Schematic Diagram of the Programming Unit

two small blurred patches on the scope. Another microswitch which completes the A.C.-circuit for the motors of the recording units closes several seconds before a displacement is shown to *S*.

(4) Recorders. It was considered desirable to have a graphic record of *S*'s movements as he manipulated the follower as well as an immediate indication of the time the follower was not on target. Two electronic units, both responding to the voltage difference obtained from the parallel network of the *S* and the program, were designed for the above uses. The circuit diagrams for these units are provided in Fig. 2.

The buffer amplifier shown in Fig. 2 serves as an isolation unit for the Brush 902 A Dual Channel Oscillograph. An isolation amplifier is necessary since it is important that the recording unit not distort or interact with the operation of any other units in the circuit. The amplifier itself should be relatively stable, *i.e.* it should not drift when the input to it is held constant nor should it respond differently to identical inputs. It is also advantageous that it be designed as a linear amplifier to facilitate the measurement of records. The input impedance of the amplifier employed in this apparatus was sufficiently high so as not to affect the other units. It also met the stability and linearity characteristics mentioned above to a satisfactory extent. The circuit for the buffer amplifier, as shown in Fig. 2, is a single stage of push-pull amplification. The output of each amplifier is applied directly to one magnetic pen unit of the Brush Oscillograph. Thus a continuous record of the location of the follower in reference to the target is obtained for each control. The total time required by *S* to reach the target can be obtained from these records. Also, a more detailed analysis of performance can be made by measuring reaction, travel, and settling times as well as by counting the number of errors and reversals. Since separate records of each control are available, it is relatively easy to compare the effect of the experimental variables upon the use of each hand.

The electronic relay is used to meet the requirement of having an immediate indication of the time *S* is off the target. If the follower is positioned over the target no potential difference exists between the leads from the control and the program to the DuMont Oscillograph (Fig. 2; Points 1 and 4, and 6 and 9). This is the condition for which the Standard Electric Chronoscopes would be stopped. When the follower is off the target a potential difference exists between those points and this voltage is used as a signal for the operation of the chronoscopes. The circuit which responds to this signal or voltage is also provided in Fig. 2. When no potential difference exists between the control grid and cathode of the input tube,

6SL7, current flows through this tube. The plate voltage of the 6SL7 provides bias for the 6SN7 of the next stage. If the plate current through the 6SL7 is of a certain magnitude then the voltage on the grid of the next tube will not be high enough to permit an amount of current sufficient to actuate the relay in its plate circuit. If a negative signal is provided to the input of the circuit, the 6SL7 conducts less current which in turn raises the grid voltage on the 6SN7. The subsequent increase in plate current through the 6SN7 may be enough to cause the relay to be energized. The input voltage necessary to activate the relay is established by adjusting the cathode resistors of the 6SL7s. Four relays, one for each half of the 6SN7s are available so that time-off target in any of the four quadrants could be recorded. For the first experiment, however, only errors in the vertical and horizontal position, regardless of direction, were desired.

The main limitation of the equipment is the inherent drift of the oscilloscope. If, for some reason, the target or follower drift from their original setting, then there will be recording errors. This means that even though *S* sees the target and follower as being coincidental on the *CRT*, there will be an error in voltage present and the recording units will respond to it. By permitting the apparatus to warm up for at least 30 min. before starting the first *S* and by checking the oscillograph between successive *S*s, the drift is sufficiently controlled. Further improvement in the over-all stability of the apparatus is desirable, however, and is being studied.

An advantage of this apparatus is its adaptability to a wide range of researches in perceptual-motor performances. With appropriate modifications it can be adapted to the study of continuous tracking.[1] Also, there is a wide range of programs that can be applied to the *CRT* without much difficulty. The use of a dual-beam oscillograph also permits the ratio between control movements and the consequent changes in the follower to be easily varied. This is accomplished by changing the amplification settings on the oscillograph. The time of the display can be controlled over wide limits and, by modulating the intensity of the program-signals, the display can be made to be intermittent.

It is also possible so to modify the apparatus that different types of control units can be investigated. At present, it is being used to study handwheel operation under various torque conditions and in various positions in reference to the body. Many other types of controls, such as stick and knob controls, could be used as well.

---

[1] Since this was written the programming unit has been modified and the apparatus is at present being used in an experiment on continuous tracking performance as a function of control crank radius, frictional loading, and position.

AN ELECTRONIC DEVICE FOR THE PRODUCTION AND MEASUREMENT OF WARBLE-TONES

By John F. Corso, Pennsylvania State University

The problem of producing and measuring the rate and the extent of a frequency-change in a tonal stimulus has often troubled the experimenter in psychoacoustics, particularly when the desired change has involved a slow rate of modulation and a minimal deviation from the mean frequency of the tone of reference.[1] While several methods are available for modulating frequency, *i.e.* producing warble-tones, none is especially adapted for the problem at hand.[2] Some methods employ fairly expensive beat-frequency oscillators; others fail to provide rates and extents of modulation which approach threshold values. Furthermore, techniques for the calibration of such auditory stimuli are not readily available. The present paper describes an inexpensive device for producing frequency-modulated tones and it outlines a specific procedure for calibrating the rates and the extents of modulation.

In the production of a frequency-modulated (vibrato, warble) tone, it is essential to have some electronic means of changing the frequency of a signal generated by a master oscillator. This oscillator may be either a commercial or a laboratory-built unit whose output approximates a sine wave. If a musically acceptable stimulus is required, a multivibrator may be employed as the generating source,[3] provided the signal is appropriately filtered to eliminate any undesired upper harmonics. In this case, a Hartley oscillator is typically used as the master unit to control the natural frequency of the multivibrator.

To modulate the frequency of the primary signal, a reactance tube (*e.g.* 6AG5) is connected in parallel with the tuned (tank) circuit of the master oscillator or, in the case of the multivibrator source, with the tuned circuit of the Hartley oscillator. The reactance tube is supplied with the usual plate voltage and cathode-bias voltage; in addition, its plate is supplied with an alternating voltage from the tank circuit of the master

[1] J. F. Corso and D. Lewis, Preferred rate and extent of the frequency vibrato, *J. Appl. Psychol.*, 34, 1950, 206-212.

[2] B. Burger, On an apparatus for the investigation of echoes in closed rooms, *Hochfrequenztechnik und Elektroakoustik*, 61, 1943, 75-82; F. V. Hunt, Apparatus and technique for reverberation measurements, *J. Acous. Soc. Amer.*, 8, 1936, 34-41.

[3] L. L. Beranek, *Acoustic Measurements*, 1949, 297-299.

oscillator through a condenser.[4] In operation, the tube develops a capacitive reactance which is instantaneously variable by means of the control voltage on the grid of the tube. Any change in the grid-voltage of the reactance tube produces a change in the frequency of the master oscillator, due to the change in reactance across the tank circuit. An increase in the negative voltage (more negative) on the grid of the reactance tube produces an increment in frequency; a decrease in the negative voltage (less negative) produces a decrement in frequency. When a sinusoidal voltage of very low frequency is applied to the grid of the reactance tube, a corresponding sinusoidal change occurs in the frequency of the master oscillator. The *extent* of modulation (limit of frequency-change) about the mean frequency of the master oscillator is, then, determined by the *amplitude* (maximal voltage-swing) of the sinusoidal voltage applied to the grid of the reactance tube. (To insure linear operation, *i.e.* direct proportionality between changes in voltage on the grid of the reactance tube and changes in frequency of the master oscillator, extreme care must be taken to set the grid-bias voltage of the reactance tube at the operating mid-point value on the tube-characteristic curve.)

The *rate* of modulation (number of pulsations per second) of the audio-signal is determined by the *frequency* of the sinusoidal voltage applied to the grid of the reactance tube. This frequency is provided by a pentode phase-shift oscillator which, acting through the reactance tube, imparts a sine-wave pulsation to the output of the master oscillator. To vary the rate of modulation (pulsation), the frequency of oscillation must be altered by introducing new constants of resistance or capacitance, or both, into the feedback circuit of the phase-shift oscillator.

Fig. 1 contains a circuit-diagram of a frequency-modulation control unit which utilizes the principles of operation described above.[5] The unit is capable of providing four predetermined rates and four predetermined extents of frequency modulation. As indicated, the rate of modulating the frequency of the master oscillator is controlled by the values of resistance and capacitance in the feedback circuit of the phase-shift oscillator (6J7). Switch 1 is a four-gang rotary switch which is used to select appropriate

---

[4] See E. C. Jordan et al., *Fundamentals of Radio*, 1942, 334-336, for a discussion of frequency-modulation and representative circuit-diagram of an oscillator and reactance tube.

[5] The control-circuits and techniques of measurement described in this paper were developed with the aid of Dr. R. H. Burkhardt, then a graduate student in the College of Engineering, State University of Iowa.

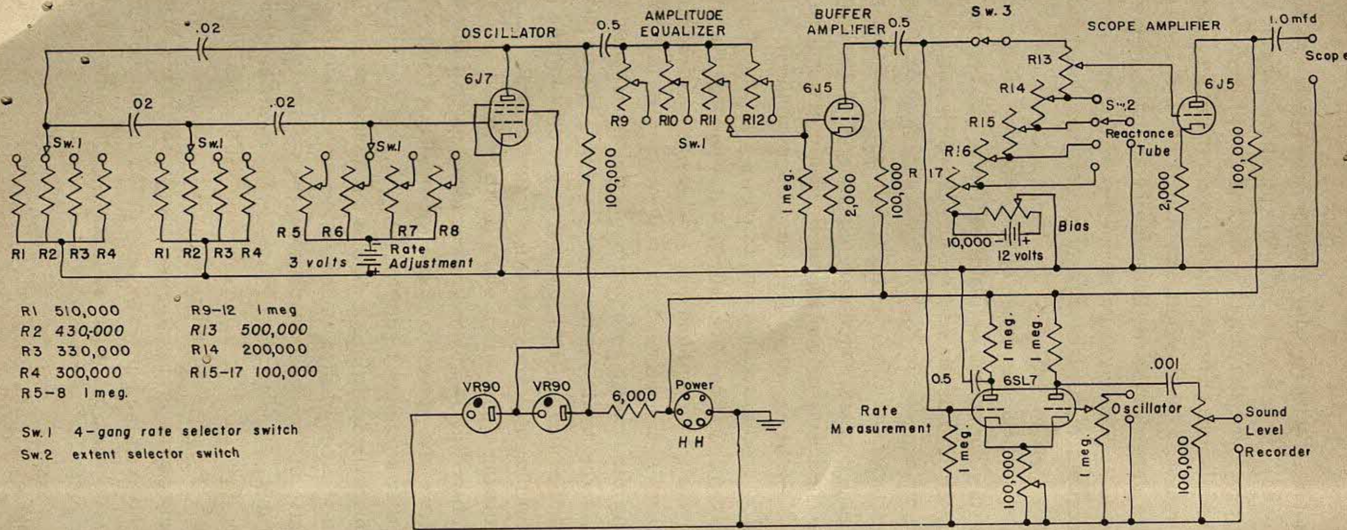


FIG. 1. CIRCUIT DIAGRAM OF AN ELECTRONIC DEVICE FOR THE PRODUCTION AND MEASUREMENT OF WARBLE-TONES

fixed resistances (sets R_1 to R_4) for a coarse rate-adjustment, a variable resistance (R_5 to R_8) for a fine rate-adjustment, and a variable resistance (R_9 to R_{12}) for equalizing the output voltage of the phase-shift oscillator at each rate of modulation. (The amplitude equalizer insures independence of modulation rate and modulation extent.) With the resistance and capacitance values shown in the oscillator feedback circuit, rates of modulation from zero to at least eight pulsations per second can be obtained. Each of the four potentiometers, R_5 to R_8, produces a different rate depending upon its exact setting.

As previously described, the extent of the frequency-change about the mean frequency of the master oscillator is controlled by the maximal voltage-swing on the grid of the reactance tube. Switch 2 is used to select from the voltage-divider (R_{14} to R_{17}) the voltage-amplitude to be applied to the grid of the reactance tube. With the values of resistance shown in the voltage-dividing network, extents of modulation (shifts in frequency above and below the mean frequency) can be obtained which range from zero to at least 4.2% of the mean frequency. Each of the four potentiometers, R_{14} to R_{17}, produces a different extent of modulation depending upon its exact setting. (The 12-v. battery and 10K-Ω potentiometer are used to adjust, initially, the optimal bias voltage on the reactance tube to insure a linear relationship between the voltage-input to the grid of the reactance tube and the frequency-output of the master oscillator.)

The function of the triode (buffer) amplifier (6J5) is to increase the output voltage of the phase-shift oscillator after equalization and to isolate the oscillator from the succeeding stages in the control unit.

A voltage-regulating circuit using two VR 90s is included to make the output of the phase-shift oscillator independent of variations in line voltage. This circuit also serves to isolate the oscillator from other units operating from the same power supply.

The calibration of the control unit used in the production of frequency-modulated tones is accomplished in the following manner. To establish the limits (extents) of frequency modulation, the voltage from the phase-shift oscillator (at R_{13}) is first passed through the triode amplifier (6J5) and presented on the screen of a cathode-ray oscilloscope. The potentiometers R_9 to R_{12} are then individually adjusted to yield readings of constant amplitude on the grid of the oscilloscope for each rate of modulation. Switch 3 between the buffer amplifier (6J5) and the voltage divider (R_{14} to R_{17}) is opened; a square wave having the same amplitude as

the low-frequency sine wave from the phase-shift oscillator, when measured on the grid of the oscilloscope, is then substituted across the voltage divider. By using a Conn chromatic stroboscope or some other frequency measuring device connected in parallel with the output of the master oscillator, the frequency limits produced by the modulating square wave are easily determined. The desired extents of modulation are obtained by making appropriate adjustments on the individual taps of the voltage divider. When each of the four desired extents has been obtained, the calibrating square wave is replaced by the original sine wave of the phase-shift oscillator. Since the extent of modulation is dependent only upon the maximal voltage-swing on the grid of the reactance tube, regardless of voltage waveform, the extents of modulation produced by the square-wave generator and the sine-wave oscillator are, in each case, equal.

To calibrate the rate of frequency-modulation, the rate-measurement circuit shown in Fig. 1 is used. This circuit (6SL7) converts the low frequencies from the phase-shift oscillator into corresponding amplitude-modulations which are impressed on a signal from an external oscillator. The amplitude-modulated signal is then fed into a graphic sound level recorder, or other paper-recording device, to provide a visual record upon which linear measurements may be made. Since the rates of amplitude-modulation and frequency-modulation are equivalent, the rate of frequency-modulation produced by the setting of a given potentiometer is determined by counting the number of recorded amplitude-pulsations in a given unit of time (linear equivalent of time). The four desired rates of modulation are obtained by adjusting individually the four potentiometers, R_5 to R_8, in the circuit of the phase-shift oscillator. (If highly precise calibration data are required, a mean value over a ten-second interval will yield a more accurate value of the rate per second than the time measurement of a single pulsation.)

To facilitate the manual presentation of modulated stimuli in an auditory experiment using, for example, the method of paired comparisons, two groups of lever-action switches may be mounted on the chassis of the control unit. One group of switches should then be connected in parallel with the rotary switch which selects the predetermined rates; the other, in parallel with the rotary switch which selects the predetermined extents. Thus, by depressing one lever in each of the two groups of switches, it is possible to combine any rate of modulation with any extent

of modulation. This provides a total of sixteen different warble-tones, all having the same mean frequency. With the addition of electronic or mechanical timing devices and appropriate filters to eliminate side-band frequencies produced in modulation,[6] the unit described can be readily utilized in many types of psychophysical research.

[6] Jordan et al, *op. cit.*, 239-240.

NOTES AND DISCUSSIONS

ONE MAN'S PREFERRED FIFTH

Recently I have been listening to three simultaneous oscillator tones in a pair of earphones and have been attempting to tune chords and measure the results. The variety of auditory (and electronic!) phenomena has been impressive. One of the simpler phenomena concerned the preferred tuning of the musical interval of a fifth, *i.e.* the distance between the top and bottom notes of the ordinary three-tone chord. There seemed to be two slightly different ways to tune this interval in the three-tone chord and obtain acceptable results. When one oscillator was dropped out and only two tones were heard, the phenomenon persisted. In the continuum of possible fifths there were two landmarks; fifths of two sizes were acceptable and good, but the larger of these was preferred. The question immediately arose: what are the sizes of the 'large' and 'small' fifths?

The apparatus set-up made it possible to measure these conveniently. The two oscillators were also connected to a cathode-ray oscilloscope, one to the horizontal plates and one to the vertical plates. The measurement procedure was to tune the interval subjectively, switch the signals to the oscilloscope, and re-adjust the frequency of one oscillator until the frequency ratio of 2:3 (a perfect fifth) was indicated by the appropriate Lissajous figure. One of the oscillators was the General Radio Type 1304-A with an incremental frequency dial. The perfect ratio was produced by changing this dial; the amount of change could then be read in cycles per second. The dial was marked in units of one cycle and was read to the nearest half cycle.

On each of two successive days I went through the following procedure. One oscillator was set successively at seven frequencies ranging from 150 to 600 ~. At each frequency I adjusted the second oscillator (higher in frequency) to produce the large (or small) fifth and then measured the change necessary to produce a perfect fifth; following this I set the higher-frequency oscillator to produce the small (or large) fifth and again measured the change for a perfect ratio. At each frequency the loudness of the tones was set at a moderately low level to minimize combination tones.

Introspectively, the small fifth was chiefly characterized by being a very smooth interval. It was also lifeless, dull, flat (in the sense that a taste is flat), and uninteresting. The large fifth sounded at least as good as an interval, and it was clear, lively, interesting, and more pleasing. As the

judgments were made, it became quite obvious that the small fifth was a perfect interval without beats, while the large fifth had beats. The beats in the large fifth did not seem to be essential aspects of the interval quality and I attempted to disregard them. It seemed to me that I was judging this fifth in terms of its quality as an interval, *i.e.* the pitch-distance from one tone to the other.

The results were highly consistent. In every case, the obtained small fifths turned out to be perfect, *i.e.* they were within one-half a cycle of being a perfect 2:3 frequency-ratio. The results obtained with the large fifths were more striking. They are shown in Table I, in which the obtained frequencies of the upper note in the fifth are shown as deviations

TABLE I

DEVIATIONS OF UPPER NOTES IN SIMULTANEOUS INTERVALS FROM PERFECT FIFTH (cycles per second)

	Frequency of lower note						
	150	200	250	300	375	450	600
Day 1	+2	+2	+2	+2	+2	+2	+2
Day 2	+2	+3	+2.5	+3	+2.5	+3	+3

from the frequency giving the perfect fifth. On the first day, every interval was two cycles per second larger than a perfect fifth; on the second day, the obtained intervals were from two to three cycles per second larger than the perfect fifth.

The results obtained on the first day were completely unexpected. I had thought that if results were uniform there would be a constant *relative* (interval ratio) difference—perhaps even the difference between a tempered scale fifth and a natural scale fifth, small though that is. The constant *absolute* difference seemed to mean that I had been making the judgments in terms of a beat phenomenon. My feeling that I had been judging in terms of interval-quality (pitch-distance) was apparently an illusion or a false interpretation.

Because I was both *E* and *O*, these things were obvious on the second day and I could hardly be called a naïve observer at that time. I could not help being aware that this was a beat phenomenon and that actually the task was to find the rate of beating that gave the most pleasing fifth. It is interesting that the most uniform results were obtained under the previous set to produce the better of two sizes of intervals.

On the second day it seemed to me that beat rates faster and slower than the preference region were definitely unpleasant. Further, the perfect fifth (actually always near-perfect), was usually changing very slowly in phase

and therefore seemed to swing around in space. This was somewhat unpleasant and disorienting. The preferred fifth, on the other hand, was beating so fast that it seemed localized and stable in space and was more pleasant for this reason.

What was the most important reason for the preference for the beating fifth? The answer probably lies in the concept of *sonance*.[1] As defined by Metfessel, this refers to the fusion in a sound of changes occurring rapidly in time, as in glides and the vibrato. The changes he noted in frequency, intensity, and spectrum were assumed to give a greater richness and pleasingness to a tone. In the interval of a fifth, beating twice per second, the chief factor in fusion must have been intensity, but phenomenally the stimulus

TABLE II

DEVIATIONS OF UPPER NOTES IN MELODIC INTERVALS FROM THE THE PERFECT FIFTH
(cycles per second)

	Frequency of lower note						
	150	200	250	300	375	450	600
Day 1	−0.5	−3.5	+4.0	+6.0	−5.5	+8.0	− 5.5
Day 2	−1.0	+4.5	+4.5	−0.5	−2.5	0.0	−13.0

was complex and there may have been fluctuation in the relative intensities of combination tones and aural harmonics.

It is interesting that the preferred rate of cyclical change in the large fifth was two beats per second, whereas in the sonance of the single tone (the vibrato) the preferred rate is six to seven cycles per second.[2] Apparently there are fundamental differences between the two cases.

If sonance was the basis of these size-of-interval preferences, my original feeling that interval quality was the criterion, or even a factor, must have been incorrect. As a check on this presumed interval quality factor, I went through the series of frequencies again, listening to the tones *alternately* and adjusting them to produce the best possible melodic interval. The results for the two days are shown in Table II. The numbers indicate how the obtained upper note differed in frequency from the perfect fifth. These are crude data because the intensity of the tones was not controlled (loudness was roughly equated). The important facts are the variability of the settings and the approximately equal number of plus and minus settings. Interval quality alone did not produce fifths consistent in either direction or size of deviation.

[1] M. Metfessel, Sonance as a form of tonal fusion, *Psychol. Rev.*, 33, 1926, 459-466.

[2] J. F. Corso and D. Lewis, Preferred rate and extent of the frequency vibrato, *J. Appl. Psychol.*, 34, 1950, 206-212.

The interval-size preference here found seems to be another example, in this case with a simple stimulus, of the point made many times in Seashore's laboratory: the most esthetically pleasing stimulus is the one in which there are slight but systematic deviations from the regular and the perfect.

University of Missouri WILLIAM H. LICHTE

A NOTE ON MONOCULAR DEPTH-PERCEPTION

The term stereopsis is commonly used to refer to the perception of depth and relief of objects, resulting from the fusion of two, slightly disparate retinal images. As such, it is a property of binocular vision. On the other hand, it has long been known that a certain degree of depth-perception is experienced monocularly. The term monocular stereopsis is sometimes used to describe this effect. In the writer's opinion, a different term should be used because binocular stereopsis produces an enhanced perception of sculpturesque relief in contradistinction to the more limited ability of the single eye to infer depth and locate successive planes in space. The term depth might well be reserved for monocular perception and stereopsis for binocular perception.

Eye specialists have long known that one-eyed individuals can judge distances in depth and even thread needles without much difficulty. Zoethout quotes Eldridge-Green to the effect that there is "undoubtedly perception of relief with only one eye."[1] He refers to the case of a man blinded in one eye who could judge distances and the size of near objects almost as accurately as when he had enjoyed binocular vision.

Secondary cues for monocular depth-perception are interposition, serial and linear perspective, relative and absolute size, relative and absolute movement, relative height, the distribution of light and shade and the kinesthetic effects of accommodation and convergence. The perception may, however, also be experienced in the form of optical illusions, when none of the secondary cues is effective or operative. Schroeder's staircase is one example.[2] It reverses itself as one looks at it into an overhanging form of masonry and the vertical plane that had appeared to be further back now seems to be nearer the *O*.

In a discussion of reversible illusions of depth, Southall wrote: "These

[1] W. D. Zoethout, *Physiological Optics*, 1935, 318.

[2] Cf. Herman von Helmholtz, *Physiological Optics*, J. P. C. Southall trans., 3, 1925, 286, Fig. 49.

illusions are purely mental and have nothing whatever to do with binocular vision as such; on the contrary, as Wheatstone remarks in his classic paper on the physiology of vision, the effects are 'far more obvious when the figures are regarded with one eye only'."[3] Southall added that true stereopsis is the result not of an optical illusion but of the fusion of two, slightly disparate retinal images.

Arnheim has pointed out that the first rudimentary indication of monocular depth-perception of an illusory nature begins to manifest itself when one looks at a circle drawn on a blank sheet of paper.[4] The plane of the circular area does not appear to be continuous with the plane of the paper but seems superimposed upon it. While there is no depth-perception as such, an *O* receives the impression that any point within the circle is

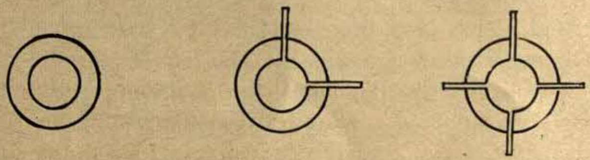


FIG. 1 FIG. 2 FIG. 3

closer to him than any point on the ground. It is not the circular area itself but the circumference that creates the effect. The latter can be intensified if the area is made darker than the ground. Conversely, we can produce an inverted, pseudoscopic effect by darkening or stippling the ground.

To produce a more definite sensation of monocular depth, however, two figures must be drawn, one within the other. If, for example, an *O* looks at two concentric circles (Fig. 1), at one time he will see a lamp-shade or megaphone, with the smaller circle nearer to him than the larger one; at another time, the figure appears reversed, with the smaller circle further away. (The circles need not be concentric, as drawn.)

It is interesting to observe that we can maintain stereopsis and prevent the figure from reversing itself by adding certain designs to the figure. If we draw two prolongations from the inner circle and make them overlap the circumference of the larger circle (Fig. 2), we succeed only partially. An *O* can still achieve the reversible effect, though with difficulty. It is only by drawing four prolongations as in Fig. 3 that we succeed completely in eliminating reversibility.

This method does not succeed with all kinds of configurations. If, for

---

[3] J. P. C. Southall, *Introduction to Physiological Optics,* 1937, 234.
[4] Rudolf Arnheim, *Art and Visual Perception,* 1954, 179.

example, an *O* gazes at Fig. 4, he will see at times a truncated pyramid, with the smaller rectangle uppermost; at other times, the pyramid seems reversed. In this case, even four prolongations from the smaller rectangle are not sufficient to prevent reversibility (Fig. 5). The reason for this is the compulsive effect of the diagonal lines, which lure the gaze into the interior of the pseudoscopic image in spite of the restraining influence of the prolongations. This is shown in Fig. 6 where the diagonal lines have been eliminated. It is now difficult, if not impossible, for an *O* to achieve reversibility.

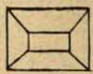


FIG. 4

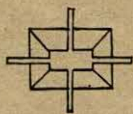


FIG. 5

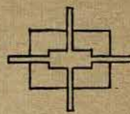


FIG. 6

These simple experiments show that it is possible for *E* to intervene in the visual-psychological conflict between stereoscopic and pseudoscopic images when an *O* stares at reversible figures and weigh the scales in favor of one or the other image by adding or subtracting cues.

A more pronounced monocular, stereoscopic effect can be produced by drawing two *eccentric* circles at the periphery of a relatively large disk presented vertically to the eye.[5] When the disk is rotated slowly, an *O* sees the lamp-shade as before, but this time from continuously changing angles of vision, so that the smaller circle seems to rotate about the center of the larger circle as both circles swing around the center of the disk. The cue to depth-perception (either stereoscopic or pseudoscopic) is continually emphasized by the changing perspective which presents different aspects of the figure to the eye, as if an actual lamp-shade were affixed at a slightly oblique angle to the revolving disk.

Three general types of depth-perception can be distinguished: (1) stereopsis based on normal, binocular vision; (2) monocular depth-perception; and (3) *illusory* monocular depth-perception. An *O* can readily note the relative weakness of an illusory depth-sensation by staring at Fig. 1 until he achieves stereoscopic depth and then quickly shifting his gaze to a pencil held in his hand. Even when viewed monocularly, the pencil shows up in much bolder relief than the 'lamp-shade' seen on the flat paper, owing to the strong contrast in luminosity between the front and sides of the pencil.

[5] M. B. Savage, The eccentric rings techinque, *Optical J. & Rev. Optom.*, 6, 1955, 35-36.

In the writer's opinion, each type of depth-perception should have a separate nomenclature, since the different types differ both in etiology and in intensity. In particular, the words depth and stereopsis should not be regarded as synonymous and used interchangeably, as they frequently are in the literature of optometry, ophthalmology, and psychology.

New York, N.Y. PHILIP POLLACK

ARE THERE DEGREES OF PREFERENCE?

I have frequently read and heard the phrase "more preferred." Is this a misuse of words? The term *preference* is derived from two Latin words, *prae* and *ferro,* and means literally to bear or place before (something else). Thus a student who goes to the movies before studying his lesson prefers, on that occasion at least, the show to his studies. There are no *degrees* of preference. *A* is before *B* or *B* is before *A,* not more or less before. If several objects are presented for choice the whole group can be arranged in a preferential sequence but still there are no degrees of preference.

Confusion comes from the fact that something does vary in degree. One object or response is more desirable, more pleasant, more acceptable, better liked, than another. There are degrees of desire, pleasantness, acceptability, appetite, liking, etc., but not degrees of preference.

Another source of confusion lies in the fact that in some experiments the subject is given a series of choices and the frequency of a preference varies. To illustrate, a rat is given repeated choices between two foods, *A* and *B.* During the first 20 choices the animal accepts *A* in 10 trials; during the second 20 choices he takes *A* in 12 trials; during the third 20 choices he takes *A* in 18 trials. The frequency of selecting *A* steadily increases with practice. In this instance it is correct to speak of changes in the *frequency* of a preference but not in *degrees* of preference.

Changes in the frequency of a preference indicate that the subject is learning a habit of preferential discrimination. The habit-strength varies with practice but this does not mean that the acceptability of the test-foods has changed.

We have found that a preferential food-habit may become so strong that it leads to inappropriate selection of foods. For example, if a rat is given a series of choices between sugar and casein, he develops a consistent preference for sugar. If he is now deprived of protein, he continues to select sugar in 'preference' to the much needed casein. When placed upon a different apparatus, however, he develops the reverse preference of

casein to sugar. The two discriminatory habits may be demonstrated in quick succession by placing the animal first on one apparatus and then on the other. True preferential relations are shown by the readiness of an animal to learn a discrimination rather than by the persistence of a preferential habit already acquired. Hence if an animal takes *A* before *B*, the preference that is apparent in behavior may indicate habit strength rather than the relative palatability of the test-foods.

Again, in the literal meaning of the word, one cannot speak of preference when a single stimulus-object is presented. In a recent experiment rats were offered a 9% solution of sucrose on odd-numbered days and a 36% solution on even-numbered days. During a 1-hr. drinking period the animals consistently ingested more grams of the 9% solution. Does this fact mean that rats prefer a 9 to a 36% sucrose solution? Certainly not, because the single stimulus method provides no opportunity for choice. Actually if a rat is given a choice between these two solutions, either in a 1-hr. drinking test or in a brief-exposure preference-test, he selects the 36% solution. Humans, also, prefer the sweeter fluid when given a choice.

The possibility of making a choice is a prerequisite of preference but the simultaneous presentation of two foods is no guarantee that a preferential discrimination will be made. It commonly happens that an animal develops a positional habit, accepting the food at the right or that at the left regardless of its quality. A positional habit precludes the possibility of a preference appearing.

In summary, the phrase 'more preferred' is incorrect because there are no degrees of preference. A preference, if it exists, can be described adequately by stating that *A* is preferred to *B* or *B* to *A*. There is no middle ground with degrees of preference. There are degrees of palatability, appetite, pleasantness, desire, and degrees of habit-strength in an acquired discrimination, but not of preference. One can state that an organism has failed to show a preference or that the frequency of a preference varied in a series of choices, but one cannot correctly speak of degrees of preference.

University of Illinois P. T. YOUNG

FIFTY-FIRST ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The fifty-first annual meeting of the Society of Experimental Psychologists was held at Brown University on April 6 and 7, 1955. Harold Schlosberg, Chairman of the Society for the year, presided at the business meeting and at the scientific sessions.

meetings on "New developments in tabulating equipment applications to tests and measurements"; "Community responsibilities of the psychologist in private practice", and "The psychometric versus the clinical approaches to management evaluation." The following invited addresses were presented: "Psychological issues in civil liberties" by Marie Jahoda; "The magic number 7 ± 2" by George Miller; and "Neocortical functions in behavior," by Karl Pribram. Films were as follows: "A long time to grow (Part 2); four- and five-year-olds in school"; "Two dynamic ambiguous figures"; and "Circling behavior in the BUE mouse strain."

Dr. B. F. Skinner presented the annual presidential address, "A case history in scientific method." During the business meeting, it was announced that the following new officers had been elected: President, Clarence Graham; Board of Directors, Francis W. Irwin and William N. Schoenfeld. Gorham Lane was reëlected Secretary for a three-year term.

The 1956 meetings of the Association will be held at the Chalfonte-Haddon Hall in Atlantic City on March 23 and 24. The 1957 meetings will be held at the Hotel Statler in New York City on April 12 and 13.

University of Delaware GORHAM LANE

SECOND INTER-AMERICAN CONGRESS OF PSYCHOLOGY

The second Inter-American Congress of Psychology was held in Mexico City, Mexico, on December 14-19, 1954, under the joint auspices of the National Department of Education and the National University of Mexico. Delegates from various Latin American countries, the United States, and Canada were in attendance. There were about 150 participants at the meetings.

The central theme of this Congress was "The Psychology of Education," which was considered from the point of view of applied psychology, of social psychology, of psychotherapy, and of teaching. Symposia were held on "Education and Applied Psychology"; "Psychology of Teaching"; and "Social-Anthropological Aspects of Education."

The program was interspersed by visits to the Mexican Mental Hygiene Clinic, Juvenile Court, Orientation Center, Medical Center for Rehabilitation of Exceptional Children, and the Psychiatric Service of the Children's Hospital.

At the business meeting of the Congress a new constitution was adopted, and it was decided to publish the transactions of this meeting, a directory of Latin American psychologists, and a bibliography of books on psychol-

ogy published in Spanish. The possibility of establishing exchange fellowships and research appointments was also discussed and tentative plans were drawn for the 1955 Congress.

The officers elected for 1955 were: President, Willard C. Olson, Professor of Education, University of Michigan; Vice-President, Guillermo Davila, National University of Mexico; Secretary, Werner Wolff, Professor of Psychology, Bard College; Treasurer, Gustave M. Gilbert. The following were elected as vice-presidents of the various hemispheral zones: *North Atlantic,* Carolina Martuscelli Bori, President of the Brazilian Association of Psychology, Sao Paulo, Brazil; *South Atlantic,* Mauricio Knobel, Minister of Education, Buenos Aires, Argentina; *Caribbean,* Jose Angel Bustamente, Professor of Medical Psychology, University of Havana, Havana, Cuba; *Central America,* Victor Funes Donaire, Minister of Education, Tegucigalpa, Honduras; *South Pacific,* Carlos Nassar, University of Chile, Santiago, Chile; *North Pacific,* Honorio Delgado, University of San Marcos, Lima, Peru; *Canada,* Robert B. Malmo, Professor of Psychology, McGill University, Montreal, Canada; *Mexico,* Eugenia Hoffs, Professor of Psychology, National University of Mexico, Mexico City, Mexico; and *United States,* Herbert S. Langfeld, Emeritus Professor of Psychology, Princeton University, Princeton, N.J.

Bard College

WERNER WOLFF

THE 1955 MEETING OF THE NATIONAL ACADEMY OF SCIENCES

The annual meeting of the National Academy of Sciences was held at its building in Washington, D.C., on April 25-27, 1955. In the morning session of April 25 there were six papers of especial interest to psychologists. L. D. Dragstedt of the University of Chicago showed that gastric ulcers are of hormonal origin due to hyperfunction of the gastric antrum, whereas duodenal ulcers are of parasympathetic origin due to hypersecretion of gastric juice. J. L. Fuller and G. A. Jacoby, Jr., of the Jackson Laboratory, reported experiments to show that food intake in hereditarily obese mice is primarily under hypothalamic control rather than sensory control. H. F. Harlow of the University of Wisconsin reported observations of learning in the infant monkey and some of the factors that motivate such learning. C. F. Richter of the Johns Hopkins Hospital described cases of sudden death in rats who, placed in a hopleless situation of having to keep swimming or drown, give up and drown. Tame white rats will persist in swim-

ming for many hours, but wild rats may give up in a few minutes. Richter drew the analogy to sudden death of human subjects in Voodoo magic, and suggested that the cause of such death may be parasympathetic and due to vagus depression, not due to hypersecretion of adrenin, as Cannon had supposed. L. L. Thurstone of the University of North Carolina described the forms of the different growth curves for seven primary mental abilities in a large school population for the ages 5 to 19 yr. H. F. Olson and H. Belar of the RCA Laboratories described and demonstrated the manner in which music can be electronically synthetized from its components by controlling the pitch, loudness, duration, rate of onset and of termination, and various other parameters of fundamental tones and the harmonics that form their timbre, as well as the accompanying noise components. The system can imitate any musical instrument or group of instruments, and provides an opportunity for new artistic creation such as no composer has ever before had at his disposal.

In the afternoon of the same day there was a symposium on sound and its perception. S. S. Stevens of Harvard University introduced the discussion with an outline of the psychophysics of hearing. R. H. Bolt of the Massachusetts Institute of Technology followed with a discussion of the physics of sound generation and propagation. G. von Békésy of Harvard University described his own experiments which show the nature of energy transmission within the inner ear. J. E. Bordley of the Johns Hopkins University discussed the aphasias where the sensory mechanisms function normally but the perception of sounds and especially of speech is blocked. Harvey Fletcher, now of Brigham Young University, summarized the symposium.

Ten psychologists were in attendance at the meetings: Boring, Carmichael, Gesell, Graham, Harlow, Köhler, Pillsbury, Richter, Stevens, and Thurstone. The deaths during 1954 of Walter S. Hunter and Calvin P. Stone were noted; also the death in 1955 of Sir Godfrey Thomson who was a Foreign Associate in psychology. The membership of the Section of Psychology now stands at twenty-four members and two Foreign Associates.

Harvard University EDWIN G. BORING

THE 1955 MEETING OF THE AMERICAN PHILOSOPHICAL SOCIETY

The American Philosophical Society held its annual meeting at the hall of the Society in Independence Square, Philadelphia, on April 21-22, 1955.

The paper of chief interest to psychologists was the report by Samuel A. Stouffer on the poll assessing tolerance of nonconformity that has now been completed under his direction. Leonard Carmichael, Wolfgang Köhler, L. L. Thurstone and the writer of this note were present. The absence of Walter S. Hunter, who died on August 3, 1954, was especially noted by his colleagues.

Harvard University

EDWIN G. BORING

ERRATA

Attention has been called to three errors which crept into my biographical sketch of Harvey A. Carr which was published in the last number of this JOURNAL (Vol. 68, March, 1955, pp. 149-151). Corrections of the errors are as follows: Carr died on June 27, 1954 (not June 21), he had an M.S. degree (not an A.M.), and he served as Chairman of the Department of Psychology at the University of Chicago in 1926-1938 (not 1922-1938).

University of Michigan

W. B. PILLSBURY

Calvin Perry Stone: 1892-1954

Calvin Perry Stone, a pioneer in comparative and physiological psychology, teacher, and scientist, died suddenly on December 28, 1954, in Palo Alto, California. Although he had suffered a severe heart attack six years earlier, he recovered sufficiently to continue his active program of teaching, research, and writing. This self-discipline and application of effort became necessary for him quite early. When he was five years old, his home burned on the day of his father's funeral, and together with his mother, sisters, and brothers, Calvin began a struggle for sustenance and education which was to last many years.

His forebears had settled in Jay County, Indiana, in 1832. Here Calvin was born on February 28, 1892. He began school at the age of six years and by fifteen had qualified for entrance into Valparaiso University from which he was graduated with the degree of Bachelor of Science three years later. His graduating class elected him vice-president, one of the year-book editors, and class poet. During the next two years he was teacher and principal of a small high-school. Summer-school work brought him to the degree of Bachelor of Arts in Classics at Valparaiso in 1913. The following year he taught in high-school. Then he entered Indiana University where, under the direction of Melvin E. Haggerty, he completed his work for the Master's degree with his thesis: "Notes on light discrimination in the dog."[1]

Haggerty continued to influence Stone's career, inviting him in 1916 to move with him to the University of Minnesota for his doctorate. While a graduate student at that University he became Director of Research in Psychology at the Indiana State Reformatory at Jeffersonville until war interrupted his graduate work in 1917 when he undertook to apply psychology to assessment and training programs of the Army until his discharge in September 1919. Thereafter he returned to his graduate work at the University of Minnesota, to be influenced primarily by Karl Lashley and A. T. Rasmussen, Professor of Anatomy. He received his doctorate under Lashley in 1921 with his thesis: "An experimental analysis of the congenital sexual behavior of the male Albino rat."[2] He then taught in the Department of Psychology and the Department of Anatomy at the University of Minnesota, impressing both students and staff with his mastery of subject matter and his expert organization of course material. Lashley said of him in 1922: "In research he has shown a good bit of originality, and a very

[1] *Op. cit., J. Comp. Psychol.,* 1, 1921, 413-431.
[2] *Op. cit., J. Comp. Psychol.,* 2, 1922, 95-153.

great perseverance showing unusual independence. His interest is chiefly in nervous and glandular mechanisms. He gives every indication that he will continue to be active in research."[3]

Thus by 1922, when Stone went to Stanford University as Assistant Professor, the course that his career was to take was clearly evident. His teaching was characterized by strict adherence to detailed subject matter presented in a formalized manner and calculated to require of the student a mastery of the facts. While his formal instruction impressed his classes, his influence was felt more by the individual students with whom he worked in the laboratory. Here his kindly interest and sympathetic understanding of a student's needs made him a helpful advisor with respect to both personal and academic problems. His research in the relationship between the nervous system and glandular mechanisms was original, a pioneering at a time in psychology when neither the physiological nor the behavioral techniques had been worked out. Yet more important was his perseverance; no investigation was complete until all facets of it had been explored. His industry, evident in twenty-three fundamental publications in his first five years at Stanford, continued throughout his life.

Though Stone's contribution to science was concerned with the relationships between physiological mechanisms and reproductive behavior, his many studies on maze learning in rats provided techniques and information fundamental to research in many areas of psychology. Stone was one of the first psychologists to recognize the great need for research in problems related to abnormal behavior and to emphasize that these problems should be of concern to physiological psychologists. His interest in abnormal behavior was expressed early in a course on this subject which he offered at Stanford in his first year there. Of particular interest is his course in 1923 on Freudian Psychology, in a sense, the first formal course on this subject offered as a regular part of the curriculum in an American university.[4] His interest in Freudian psychology continued throughout his life as indicated by his reviewing the *Sigmund Freud Letters* during the week of his death. He contributed significantly to the research in the area of abnormal

---

[3] Personal communication from Lashley to Lewis M. Terman.

[4] It is a question as to whether G. Stanley Hall's announced instruction in psychoanalysis in 1916-1920 included what can be called a formal course. In the catalogue he said that psychoanalysis is part of the "complete course in psychology at Clark University," and he mentioned Freud, Jung, Janet, and Adler by name. No special course in psychoanalysis was listed under Hall's name until 1920, the year before his retirement, and that need not mean that Hall, even in 1920, limited himself to any particular topic in any particular hour. (This last comment is furnished me by E. G. Boring.)

behavior particularly with respect to electroconvulsive shock. His interest in psychological research with human subjects, which was the concern of his first research paper "What a pupil should know before being permitted to enter high school,"[5] led Stone to participate in research on human subjects whenever the opportunity presented itself. In 1928 he spent his sabbatical year at the Institute for Juvenile Research in Chicago, and in 1945 at the Psychiatric Institute in New York; on many separate occasions he joined groups at Stanford to advance research on various aspects of behavior as related to physiology.

Calvin Stone contributed considerably to the profession of psychology. By example he did much to demonstrate the usefulness of psychology to the military service in World War I. In 1932 he toured Europe, visiting centers of interest to psychologists, so that he might approach psychology with a broader perspective. He was extremely active in various professional societies, holding office in the American Association for Advancement of Science (vice-president for Section I, 1938-39); Western Psychological Association (president, 1931-32); American Psychological Association (president, 1941-42). In addition he was an active member of the American Association of University Professors, American Association on Mental Deficiency, California Academy of Science, Western Naturalists, Society of Experimental Biology, National Academy of Science, and Sigma Xi. He was also Editor of *Journal of Comparative and Physiological Psychology* in 1947-50, Editor of the *Annual Review of Psychology* from its inception in 1950, Editor of *Comparative Psychology, 3rd edition, 1951,* and he contributed chapters to many books on the physiological mechanisms of behavior.

Stone was uncompromising in his insistence that a psychologist must be well-grounded in fundamentals. When many professors were following the pedagogical trends and becoming less formal and less exacting, he continued his characteristically formal and exacting courses. He would have preferred that all the courses in psychology be presentations of facts, presented with rigorous respect for 'evidence.' This is not to say that the content of his courses did not change; on the contrary, his courses in physiological, comparative, and abnormal psychology were being continually revised to report the most recent findings. His uncompromising attitude in respect of fact and evidence was reflected in his insistence on retaining German as a foreign language requirement for psychologists, and in his maintenance of high standards on examinations and theses. His rigor also carried over

[5] A paper presented to the County Institute, a teacher's professional society, in 1911.

into his editorial work, becoming evident in the careful reporting that he required of contributing authors. Nor was his insistence on rigor, perseverance, and industry less in the laboratory. Sometimes these attitudes were, indeed, distressing to his students, but Stone, aware of this fact, nevertheless felt his exacting approach must be maintained if psychology were to remain a science.

Calvin Perry Stone's devotion to psychology as a science, to his teaching and research, will continue to be recognized in his many writings and reflected in the work of his students. His quiet kindness and sincere helpfulness will be remembered by those many colleagues and students who felt the warmth of his friendship.[6]

National Institute of Mental Health
Bethesda, Maryland

H. ENGER ROSVOLD

---

[6] The writer is indebted to Mrs. Stone for much of the material of this biographical sketch, which, because of its professional orientation, omits a very important portion of Calvin Stone's life, his relation to his family.

BOOK REVIEWS

Edited by M. E. Bitterman, University of California

Art and Visual Perception: A Psychology of the Creative Eye. By Rudolf Arnheim. Berkeley and Los Angeles, University of California Press, 1954. Pp. x, 408.

This is a fascinating, unique, and important book. It is the first systematic application of our knowledge of perception to problems of the visual arts, and it contains much that is new to psychologists, estheticians, artists, art-historians, and art-educators.

The opening chapter on *Balance* introduces the reader to the concept of perceptual forces as he discovers the effects of changing the position of a black disk on a white square. A second disk is added and leads to a general discussion of pictorial balance. From symmetrical and asymmetrical balances of areas and their subdivisions, we go to the balance of lines and to those factors which determine balance: weight and direction, along with the position of the figure within the larger frame. Examples from pictorial art demonstrate axes of pictorial composition and lead to an examination of patterns of balance and the inhomogeneity of visual space. Since balance, counterbalance, and imbalance convey impressions of rest and activity, there arises the problem of the communication of meaning. The whole chapter is brought together in an analysis of a single work of art, *Madame Cezanne in a yellow chair.* Throughout the chapter, as throughout the book, the phenomena discussed are demonstrated by means of many drawings and pictures.

The second chapter deals with *Shape, i.e.* with the features which articulate the characteristics of objects. Shape is essentially the spatial definition of the object and may or may not coincide with its geometrical boundaries. The example is given of a person describing a circular winding staircase with a circular motion of his finger in the air: he has described the shape of the staircase. The influence of past experience is acknowledged, along with that of set and personal need, provided the structural characteristics are relatively weak and ambiguous. The fundamental organizing principle of shape is that of perceptual simplicity, which may be defined "by the number of structural features that make up a pattern. In an absolute sense, a thing is simple when it consists of a small number of structural features. In a relative sense, a thing has simplicity when it organizes complex material with the smallest possible number of structural features" (p. 41). Rubens' artistic productions may be seen as an example of the latter type of simplicity, paintings which "organize a wealth of meaning and form in an overall structure that clearly defines the place and function of every detail in the whole" (p. 38). Ultimately, perceptual simplicity is to be accounted for in terms of "a tendency to simplest structure in the brain-field," the former being the isomorphic counterpart of the latter. Admittedly, however, perceptual simplicity also depends on the stimulus, traces, needs, and sets. Perceptual veridity is guaranteed because nature also grows according to a principle of simplicity, and man-made objects conform because they themselves are the result of human perceptual organization. Perceptual similarity is the single most important way to perceptual simplicity, and Wertheimer's laws of organization can be shown to be applications of the similarity principle to the various aspects of the stimulus. These laws not only make for perceptual grouping, but they also support symbolic

meaning in art. Visual counterpoint is attained by pitting the various aspects of similarity (*e.g.* shape, position, motion) against each other. The chapter closes with a discussion of "The Structural Skeleton," and a demonstration of how the rotation of a triangle establishes the identity of a pattern.

The next chapter concerns *Form,* the content conveyed by the shaped object which is, however, not to be confused with its subject matter. The problem is introduced with a discussion of the effects of changes in the orientation of geometrically constant shapes. Orientation is always relative to a framework which includes the structural skeleton of the surrounding visual world, the brain-field on which the image is projected, and the structural framework of *O*'s body as determined by static and kinesthetic factors. There follows an analysis of the projection of two- and three-dimensional objects, and the conditions for shape-constancy. Here the reader learns how to distinguish a Mexican wearing a large sombrero from a doughnut, and what can make the distinction so difficult. Perhaps even more instructive are the twenty-seven schemata from children's reproductions from memory of "a three dimensional picture of a chair drawn in correct perspective." The problem of two-dimensional representation of three-dimensional form has had to be faced by all painters, and the author analyzes various solutions. A so-called "primitive" solution is that of the old Egyptians which turns out to be no less ingenuous than that of later perspective. Both procedures let a particular aspect of the object represent the whole, and both are successful when they indicate the incompleteness of the drawing, and when the whole structure which the incomplete drawing suggests corresponds to the correct or intended one. Foreshortening and overlapping are means to this end, and their special contributions are illustrated in detail. This discussion leads into the problem of depth in pictorial representation, which is taken up again later in the chapter on *Space.* Meanwhile there is a brief discussion of the dynamics of obliqueness, of what has been called realism in art, and of equivalent solutions in modern art. Transparency, the product and complement of overlapping, not only makes for perceptually simple and complete figures, but introduces pictorial depth. A short but most delightful section on "Competing Aspects" deals with the simultaneous representation of what are usually considered incompatible aspects and leads the author to reject the traditional theory of empathy—a rejection which is given further support throughout the book.

Another deviation of artistic form from "realism" is that of the more or less extreme limitation of the representation to very few lines or brush-strokes, a technique that is successful when the essential structural features are reproduced. The reasons for such reduced representation may be several. Children's drawings exhibit this characteristic, as does some so-called primitive art. It is also seen in some modern work, where it may be an expression of the artist's withdrawal from the here and now, as shown for instance in non-objective painting. In this connection the author discusses artistic invention and imagination, concluding the chapter with a short section on visual information.

The chapter on *Growth* is based on detailed analyses of many drawings and paintings produced by children at different age-levels. Evidence is presented which leads to the rejection of most traditional theories concerned with children's artistic productions. The old standby that "the child draws what he knows, rather than what he sees" is refuted systematically. In the first place, the child is relatively

lacking in knowledge and therefore depends more than the adult on what he perceives, with verbal knowledge entering only when the prceptual features are undistinguished. That the child does not lack the necessary perceptual acuity is evidenced by the fact that he distinguishes different persons with ease, and even different facial expressions of the same person, at a time when his drawings are still quite undifferentiated. Nor can lack of motor skill be held responsible, even though it does present some interesting problems. The essential difficulty must lie therefore in the process of representation itself. The child grasps the perceptual structure, *e.g.* the roundness of the head, but he has to invent a method for reproducing it. Making a circle on paper to stand for the roundness of the head is such an invention. In order to explain this process of artistic invention, Dr. Arnheim introduces a discussion of the medium and of representational concepts:

"Representation never produces a replica of the object but its structural equivalent in a given medium. Apart from other reasons, this is true because replication is possible only if the object is duplicated in its own medium. Anywhere else there are considerable differences between model and image. . . . The young child spontaneously discovers and accepts the fact that a visual object on paper can stand for an enormously different one in nature, provided it is its structural equivalent in the given medium. . . . The psychological reason for this striking phenomenon would seem to be, in the first place, that in human perceiving and thinking similarity is not based on piecemeal identity but on the correspondence of essential structural features; secondly, that an unspoiled mind spontaneously understands any given object according to the laws of its context. . . . Representational concepts, that is, the conception of the form by which the perceived structure of the object can be represented . . . find their external manifestation in the work of the pencil, the brush, the chisel" (pp. 132-133).

Important developments take place, however, prior to the emergence of the first representational concepts. As gestures may describe the shape of objects by their outlines, so the early drawing of outlines may be little more than recorded gestures. The first scribbles of the child probably have nothing to do with representation, but exhibit the child's discovery of bringing about something that was not there before. The shape, range, and orientation of these scribbles are largely determined by the mechanical construction of arm and hand, as well as by the child's mood and temperament, and by the medium at his disposal. The "primordial circle," developing out of the earliest circular scribbling, indicates organization according to the principle of simplicity in motor behavior, which in turn produces visually favored simplicity of circular shape. Along with these two aspects there is a third, especially intriguing one: the motor act itself is *part* of the child's representation, and only gradually sinks to the rôle of *means* of representation. Other characteristics of the circle continue to favor it after representational concepts begin to evolve. It is the earliest form of "thingness," of "containing," and as such it accounts for "the misnamed tadpoles," the *hommes têtards* which Dr. Arnheim subjects to a systematic analysis and reinterpretation. In this connection the lack of the third dimension in children's art is discussed, and illustrations from the history of art are presented to show various forms of attack on this problem. Finally we are let to some serious consideration of the educational consequences of what we have learned in the chapter:

"Misinterpretations are inevitable if the picture is considered a more os less correct replica or derivative rather than a structural equivalent of the object in terms of the medium. This holds true not only for the work of children, but for all art,

including the realistic. . . . Pictorial form develops organically from the simplest to increasingly complex patterns. Thus the process of growth gives further evidence of the tendency to simplicity which was demonstrated earlier in visual organization. Step by step the maturing mind requires greater complexity, but the higher stage can be reached only by way of the lower ones. The mastering of a given stage creates need and readiness for the next" (p. 165).

The author pleads for an integration of the cognitive and the need-projective approaches in art-education, which has adopted too one-sided an emphasis on motivational and personality factors in art work. A section on three-dimensional representation in sculpture and its special problems concludes this most interesting and rewarding chapter.

Next we approach the problem of *Space*. If a line is drawn on a white surface, the latter is not seen as interrupted by the line, but rather as continuing behind the line. Thus two planes are established, each complete in itself, and making for a simpler overall structure than a single interrupted plane would permit. The same phenomenon is observed if an outline figure is substituted for the line. As long as a two-plane organization pertains, the contour belongs to one plane only, ordinarily to the smaller plane, utilizing the convex aspect of the outline as the contour. Physiological factors make for additional differences between the space enclosed by the outline, the 'figure,' and the enclosing space, the 'ground,' with the former looking more solid, more dense than the latter, and resisting disruption more successfully. From this rather classical introduction to the figure-ground problem in terms of a two-plane relationship, the author proceeds to a more complex situation which is more typical of our ordinary visual world and of painting. Utilizing a four-plane pattern, a wood-cut by Arp, he demonstrates how alternative figure-ground organizations alter the perceptual structure of these more elaborate stimulus-conditions. Painters must make such analyses—consciously or intuitively—in order to obtain the desired depth-relationships among the various parts of the picture. For sculpture, the figure-ground problem is of similar importance, though discussed here primarily in its application to convexity-concavity aspects. The statue and the surrounding space can be considered as two volumes, with the statue monopolizing the figural characteristics of having smaller volume, being enclosed, having density, surface-texture, solidity, and, in classical sculpture, convexity, although there are examples throughout the history of art of concave boundaries, culminating perhaps, in their more frequent use in modern sculpture. Examples from Henry Moore illustrate "the admission of empty volume as a legitimate element of sculpture," and the author speculates "that this daring extension of the sculptural universe may have been made possible by an era in which flying has taught us through vivid kinesthetic experience that air is a material substance like earth, or wood, or stone, a medium that not only carries heavy bodies, but pushes them hard and can be bumped into like a rock" (p. 195).

The problem of overlapping which was discussed in earlier chapters is taken up again at this juncture, and the conditions for depth through overlapping are spelled out in detail. The spatial effect depends on the interruption of contours. Mere touching will usually not produce a depth effect, and crossing without interruption will at most result in a weak impression of distance. When there is interruption of contours, that figure whose contours have not been interrupted will be seen to lie in front of the figure whose outline has been broken. Another problem previously

mentioned in this connection is that of distortion. Greater simplicity can be achieved if an irregular stimulus-figure is perceived as a simpler, more regular figure in an oblique position than if it is seen as an irregular figure in the frontal plane. The impression of distance is thus created through systematic distortions of perceptually simple figures. Dr. Arnheim tries to account for these effects in terms of a neurological theory which, much like certain perceptual phenomena, appears to be more simple than accurate. The fact that it is simple recommends it, while its questionable veridity probably does not distinguish it from most other neurological theories.

The question of the truthfulness of spatial organization is, in fact, the next problem to be dealt with. A number of illusions are discussed here, especially architectural, three-dimensional illusions, less well known to psychologists than the Müller-Lyer figures, and therefore all the more intriguing. As in all similar discussions the final question is not: why are we fooled? but rather: why are we not fooled much more often? Part of the answer has previously been given. Classical theories explained our relatively well-adapted behavior in terms of past learning, based not only on visual but also on tactile and kinesthetic cues. On the whole, these explanations are convincingly rejected.

Objects influence the character of the space around them, *i.e.* the spatial framework itself is a function of the perceived objects. Alternate ways of viewing a trapezium demonstrate once again organization of space according to the principle of maximum simplicity. The fact that distortions are only approximately compensated leads to a consideration of some much neglected characteristics of perceptual space—its inhomogeneities and its non-Euclidian character—which can be derived from the range of facts which have been discovered in experiments on size-constancy. Gibson's gradients are considered in this connection, and their manifestation in oblique representation. The remainder of the chapter shows how these various principles are applied by the artist for the creation of desired spatial effects, denoting both perceptual and "symbolic" space.

The chapter on *Light* opens with an analysis of the phenomenal experience of light and darkness. Brightness-constancy is considered as the outcome of the perceptual separation of illumination and object-color. The problems of illumination, transparency, and the spatial effects of light and shadows are elaborated in some detail before the author proceeds to their utilization by the painter, again both for realistic and symbolic effects.

The effects of *Color* have been discussed by many writers in different fields. In a critical consideration of "warm" and "cold" as attributes of hues, the author contends that these characteristics may be more pronounced in mixtures than in pure colors. Most of the chapter deals with questions of color-harmony, color-scales, color-mixture, and complementarity, culminating in the analysis of color-effects in two paintings: Matisse's *Luxury* and El Greco's *The Virgin with Santa Ines and Santa Tecla.*

The last three chapters are concerned with problems of pictorial "dynamics," *i.e.* movement, tension, and expression. It must be emphasized, however, that this is not their first appearance. Throughout the earlier chapters they are referred to repeatedly, and detailed discussions of special relevant aspects can be found throughout the book.

Movement attracts attention more than any other perceptual attribute of an

object. Movement is an event and yet, with the exception of *phi,* we do not see events, but objects undergoing change in time.

"Time is the dimension of change. It helps describe change, and does not exist without it. In a universe in which all action had ceased, there would be no more time. Similarly immobile objects give us the impression of being outside of time. Theoretically my desk persists in time while my pen is moving over the paper. But I do not perceive the desk as being busy persisting the way the pen is busy moving. At any given moment the pen is in a particular phase of its course across the paper. For the desk there is no such comparison between its states at different moments. It does not 'remain the same' or 'stand still.' The time dimension does not apply" (p. 305).

Yet, there is an experiential timelessness of many events, such as the leap of a dancer, or a car which is seen moving through space rather than through time. Time is therefore a functionally distinguishing feature of motion rather than its perceptually distinguishing feature. The perceptual experience of movement rests on the organization of a sequence of events, rather than on the integration of simultaneously presented data. This interpretation implies the continuous availability of memory-traces. It also prescribes a time-ordering of the perceptual-mnemonic process. The order of percepts is not part of the composition of a picture or sculpture when we look at it, or at least to a far lesser degree than it is for a dance. From these considerations the author proceeds to the conditions of perceived movement. Two systems must be seen as being displaced with respect to each other. Other things equal, the 'figure' is seen as moving against a stationary 'ground.' If the framework is seen as moving, then any dependent figure at rest is perceived as either incapable of, or actively resisting, motion. Special problems arise with respect to the path of a moving object. Duncker's example of rotation- and translation-movements of the wheel are supplemented by others, and the author concludes that in any but the simplest movements a separation of the over all pattern occurs that makes for the perception of several simpler systems rather than the complex, undivided whole-movement. There follows an interesting analysis of perceived speed of movement as it is affected by, and in turn affects, the perception of both figure and ground. Stroboscopic movement and its application to the cutting and editing of film are considered in some detail. Michotte's interesting experiments, which have been all too slow in finding their way into the American literature, are presented in this connection. The differences between the visual and the kinesthetic body-image are discussed in terms of the kinesthetic experience of the dancer and the visual experience of the onlooker, and some of the unique features of motion pictures are commented upon.

Tension refers to experienced motion without experienced displacement—to all motion conveyed in pictures or statues, as contrasted with the dance. Although lacking displacement, tension has direction, and is literally defined in terms of direction. Here, as previously, the author makes a strong argument against a kinesthetic-empathic theory. The remainder of this chapter is concerned essentially with ways in which the artist can convey tension, which is largely in terms of the factors discussed in previous chapters.

In the last chapter, on *Expression,* various theories concerning the expression of emotion are presented. The author argues against expressions as 'learned stereotypes' and for the priority of physiognomic experience. The implications of this problem

for art-education are pointed out and the book concludes with a few brief comments on symbolism in art.

This summary of Arnheim's book is not only incomplete as a general outline, but does not sample the choice morsels which make the reading so fascinating, enjoyable, and instructive. Much new material could not even be mentioned for lack of space. Dr. Arnheim presents a systematic theory of perception which brings together the facts of experimental psychology and artistic production. The theory is firmly rooted in the tradition of Gestalt psychology. Wertheimer's laws of perceptual organization, Koffka's discussions, and Köhler's isomorphism provide the *sine qua non,* but the author transcends the limits of the classical treatment. It might be well to note that there is no attempt to deal with the recently popularized relationship between perception and motivation and between perception and personality. While recognizing that artistic productions offer rich material for those topics, the author is concerned here with universal aspects rather than with those relevant to the study of individual differences.

The book would have merited much more careful attention from its editors. The general organization could probably have been improved somewhat, and occasional less fortunate expressions could have been eliminated. More disturbing is the 'artistic' make-up. As the reader opens the book he stares at five inches of white space in the center between opposing pages, while the text on these pages looks as if caught in the act of eloping over the upper and outer margins. One might recommend that the designer read the book, and this reviewer stands ready to supply relevant page-references.

Finally, one should mention some of the things this book has lots of. First, illustrations—there is hardly a page without at least one drawing, and there are a good many reproductions of works of art. Then there are a large number of references, many of them to less well known works. There are 15 pages of annotations to the chapters and a carefully prepared index. Last, but perhaps most important, there are problems. Throughout the book the author points to problems which should be investigated, and to hypotheses which await experimental support.

College of Medicine, University of Illinois

MARIANNE L. SIMMEL

Areas of Psychology. Edited by F. L. MARCUSE. New York, Harper and Brothers, 1954. Pp. viii, 532.

Reviewing an assemblage of chapters written by authorities in their respective fields is a delicate, indeed, a rather presumptious task for one person. In the present case, the reviewer has been impressed with the scholarly tasks accomplished by the thirteen contributors, and the accompanying remarks will bear principally upon the usefulness of this volume for instructional purposes at the undergraduate level for which it is intended. A general evaluation can best be supported by brief summaries of the respective chapters which represent, as the title indicates, the major areas in which psychologists are currently working. Unfortunately, Marcuse has provided no biographical data for the respective contributors.

Nelson's chapter on Vocational Guidance presents a comprehensive and lucid review of the field. Types and sources of data needed for vocational guidance are described, and arguments are presented for a more directive approach to the coun-

seling interview. A notable feature of the chapter is the section on problems of evaluation in vocational guidance; two sample studies are described. In concluding his discussion, the author lists current standards for vocational guidance service suggested by the National Vocational Guidance Association. This emphasis on current concern with professional standards in applied areas is echoed in subsequent chapters of the book and is a noteworthy addition to the usual textbook material.

Industrial Psychology is handled in two chapters. The first of these, by Patricia Smith, deals with the Selection and Placement of Workers. The author defines in detail the rôle of the industrial psychologist in this phase of industrial planning. Problems in the establishment of selection procedures are made meaningful by use of an illustrative problem, around which most of the chapter is organized. Critical attention is given to the interview as a diagnostic instrument. The tenor of this discussion is indicated by Smith's statement that "Unfortunately, objective studies . . . show that the performance of most interviewers is characterized much more by self-confidence and enthusiasm than by accuracy." Evidence is cited in support of this contention, and suggestions are made for more successful use of the interview. Ryan's chapter on Efficiency in Work as an aspect of industrial psychology is not as well organized as its running mate. Early in the chapter, the author states: "We shall consider the term (total) efficiency to mean the ratio between what the worker accomplishes and what he puts into the work. . . ." Unfortunately, this promise of a quantitative relationship between effort and achievement is not kept in the ensuing discussion. Factors comprising the denominator of this ratio turn out to be often tentative and frequently unmeasurable. At least one highly debatable assumption is presented, namely, that "motivation and effort are closely related terms." Ryan's attempt to clarify this relationship by making a distinction between *intrinsic* and *extrinsic* motivation seems only to make matters worse and leads to the puzzling assertion that "monotony or boredom is a state of negative motivation." In contrast with the two preceding chapters, the rôle of the psychologist in this area is not well defined. In motion study, for example, where does the industrial engineer leave off and the psychologist begin? The author concludes by pointing out appropriately how little we actually know about the effects of such specific factors as noise, temperature, and lighting, upon efficiency.

Legal psychology is represented in rather attenuated fashion by Weld's chapter on The Psychology of Testimony. Despite its many archaic features, this section will interest the student, who is shown the factors contributing to the fallibility of courtroom evidence and the effects of testimony upon jury decisions. The discussion of experimental studies of perception, however, is woefully out-dated. Evidence is drawn largely from protocols derived from Leutemann's colored picture of a scene in India. Subsequent report is described in terms of the classical *Aussage* experiments, with no mention being made of the recent work of Allport and Postman on the transmision of rumor. The author also makes several broad generalizations unsupported by reference to data, such as: "There seems to be no question but that men are on the whole less susceptible to suggestion than women. . . ." The reviewer would have been less troubled by the assertion that "the testimony of witnesses of high moral character such as clergymen, police officers in general, lawyers of high standing, etc., has great weight" if he were certain that the author had written it with tongue in cheek.

Corsini has done an excellent job of describing Criminal and Correctional Psychology. The author takes a psychological viewpoint with respect to crime, contrasting it with sociological and other positions. Topics covered include theories of criminal behavior, the functions of the correctional psychologist, and research in criminology. Both individual and group therapy are outlined. The latter is documented with several case histories and excerpts from one actual therapeutic session, replete with some rather 'salty' language that is certain to delight the undergraduate reader. Corsini is critical of the present penal system, as are most authorities, and is frank to admit problems in the etiology of criminal behavior that remain unsolved.

Social Psychology is represented through the eyes of MacLeod, who opens his discussion with a forthright attack upon some conceptual problems involved in the behavioral and phenomenological points of view. The reviewer waxed alternately hot and cold in reading this chapter. MacLeod's opening statement on the province of social psychology, as compared with general psychology and sociology, is one of the clearest available, but his treatment of "four types of social psychology" produces some strange bedfellows—the group-mind notion in company with theories of social motivation, learning, and group-dynamics. Social motivation is discussed in terms of needs, goals, and regulators of behavior. From a relatively simple and straightforward beginning, however the author launches into a theoretical discussion, which the reviewer feels to be rather too involved for comprehension by the average undergraduate with one previous course in psychology. For example: "The principle of functional autonomy asserts that, however, a need-goal relationship may have developed, it may become established in the psychological field as an identifiable, relatively autonomous system." We learn shortly, however, that "some motivational systems are clearly less autonomous than others," and that "human motivation may be thought of as a complex hierarchical organization of interrelated systems." The reviewer will certainly not dispute these contentions, but he doubts seriously that sophomores will entertain them any longer than necessary to pass the course. The author descends from these flights of oratory to deal ably with the notion of "psychological groups," including national and racial group-memberships. A final section is concerned with communication, persuasion, and leadership.

Murphy and Rabban present the field of Child Psychology as an important human concern. Many of the traditional topics are discussed, *i.e.* methods of studying child behavior, maturation and learning, intellectual development, culture and personality. Aggression, competition, and reactions to authority are discussed in some detail. A particularly interesting section of this chapter deals with "problem children and children's problems," in which the authors point out that the objective situation is important in the adjustment of the child principally through the child's perception of the situation. In the related chapter on Educational Psychology, Freeman provides an orientation for the student by tracing the history of educational theory. The author not only describes principles in the personality development of pupils, but also mentions personal characteristics desirable in teachers. A large part of the chapter is devoted to the topic of individual differences. The material is recent and the discussion engaging. In the section on learning theory as applied to education, a few unfamiliar notions appear. Several of the approaches to learning are described as conditioning, mechanical repetition, trial and chance, retrial for motor skill, insight through visual survey, and reflective thinking. This analysis seems over-segmentalized

and not entirely in harmony with concepts of learning that the student will probably have acquired in a general course. On the other hand, the student interested in educational psychology may as well be exposed to some of the jargon.

The introduction to Abnormal Psychology by Mittlemann and Malkenson is too brief to provide an adequate orientation and will serve mainly to frighten the student with unfamiliar names and concepts. The entire coverage of this area, while thorough, is necessarily highly condensed and may prove difficult for the student to grasp without considerable supplementation from classroom lectures. Several weaknesses are apparent. First, no distinction is made between therapies useful in the neurotic as distinguished from the psychotic disturbances. Secondly, the discussion of scoring and interpretation of Rorschach and *TAT* protocols seems singularly inappropriate at this level of instruction. Otherwise, the chapter is competent and well-organized. The following chapter on Clinical Psychology, by Steinzor, is geared more closely to the needs and preparation of the intended audience. A unique and commendable feature of this discussion is the attention given to the development and current status of clinical psychology as a profession. The author describes penetratingly the problem of therapy in terms of the clinician-patient relationship. Throughout the chapter, the importance of research in clinical psychology is emphasized. Training of the clinician, legal and ethical problems, and professional conduct are other noteworthy features of Steinzor's summary of this area.

Riess discusses at some length the underlying philosophical assumptions of Physiological Psychology. Fundamental biochemical concepts are presented briefly but clearly. The author's review of recent trends in this field provides the student with a perspective on research possibilities, although the material becomes somewhat overly-technical when Hebb's cell-assembly theory and Selye's concept of the general adaptation syndrome are introduced. Birch describes the evolution of Comparative Psychology as due to "efforts to examine the emergence of psychological capacity, of behavioral mechanisms for coping with the demands of life." He discusses anecdotalism and the law of parsimony, as well as the problem of instinctive behavior. Learning experiments with different species are dealt with under the more general heading of "modifiability" of behavior. The relevance of comparative studies for understanding human behavior is pointed up in a section on the evolution of social life. The concluding paragraphs on abnormal behavior in animals are too brief to be effective, and no illustrative experiments are cited.

The final chapter of the book, on Psychology and Esthetics, by Gundlach, is difficult to evaluate. The criticism of over-compartmentalizing which might be leveled at the book as a whole is certainly justified by treatment of esthetics as an area of psychology. Justification for including esthetics as a separate topic is further weakened by the tone of the writing, which seems interpretive and polemical rather than factual and objective. The author begins on a querulous note by deploring what he feels to be a lack of sensitivity toward artistic products on the part of the average American. The ensuing discussion seems only vaguely related to psychological principles which, although enunciated, do not appear critical to the author's analysis. In short, there is nothing peculiarly 'psychological' about the chapter; it could as well have been written by a philosopher or some other astute critic of the arts. The style of the chapter, however, is engaging, and the content is novel compared with the more traditional topics in psychology.

Having taught courses in both general and applied psychology, the reviewer has faced several times the problem of textbook selection. The decision has seldom been easy. In general, the most effective texts appear to be those written by one or two authors. Such books have a uniformity of conceptualization and an evenness of treatment that is seldom achieved in volumes compiled under multiple authorship. While less stimulating to the professional reader, the single-author approach is undeniably better suited to the abilities of the average student. The present book seems to be no exception to this general rule. The editor, of course, would like us to believe that the book is suitable for non-majors in psychology, who have had only an introductory course as preparation, as well as for majors who want some idea of problems and methods in the different areas of psychology. The fact is that in many departments a course in Fields of Psychology will be largely repetitious in the case of majors, who will cover the same material in more advanced courses. Non-majors, on the other hand, do not ordinarily have the factual background to fully understand and appreciate brief, condensed treatments of the several areas of specialization within psychology. As indicated, the book under consideration is somewhat uneven in the degree to which it provides for the latter type of student.

These shortcomings, of course, hold true for any book of this type. Among the volumes available for a terminal course in psychology of the 'fields' or 'applied' variety, the present book merits serious consideration. Its defects are more than balanced by its generally recent and scholarly coverage of the more distinguishable areas within psychology.

University of Maryland — ELLIOTT McGINNIES

Dvorine Pseudo-Isochromatic Plates: Twenty-Three Plates, with Instructions, Nomenclature Color Disk, and Sample Record Blank. By ISRAEL DVORINE. Second edition. Waverly Press, Baltimore, 1953.

Announced on the title page as a "second edition" but copyrighted (on an inner one) as a "revised edition" of the ponderous two-volume *Color Perception Testing and Training Charts* of 1944, this work is really a far cry from the earlier one, with only two or three of its 130 plates retained, and most of its unique and semi-heretical assumptions wisely jettisoned. A compact 7 × 6 in. booklet in vivid indigo cover (encased in black), it contains 23 plates, one to a page, in the stock makeup of its Ishihara model: a 3½ in. circle of colored dots or bubbles on a white ground, mounted on thin black cardboard.

Section One consists of 15 plates. One carries a demonstration-pattern in red and blue; the remaining 14 show a two-digit (sometimes a one-digit) number in one color set askew on a ground of a second color, with two plates or examples of each of the 7 pairs of color utilized. The training program of 1944 (probably generated by war pressures) has passed into the discard. The pairs of colors chosen are not haphazard combinations as formerly, but selected as genuinely pseudo-isochromatic for the red-green defective (for whom alone the test is planned). The designer, an optometrist, acknowledges helpful criticisms by Louise Sloan of the aptness or inutility of all but 11 of the combinations of the 1944 edition; suggestions also from Deane B. Judd relative to the 7 pairs of colors used in the new text. Section Two, designed for children, for illiterates, and for corroborative use, features the trail introduced by Ishihara, with one example each for the seven

pairs of colors used in Section One, plus a red and blue demonstration-pattern.

As to general makeup of the test, instruction, aim, and scoring, certain improvements over the 1944 schemata may be noted. As before, an identical patterning of dots on all the plates—a device now generally accepted as both economical and as minimizing the use by color-deficient cases of casual clues to identify a plate—is adopted. A loose-leaf binding now permits an altered order of the plates as a further deterrent to memorizing (rendered difficult also by the use of two-digit numbers). The *instructions for administering* the test are more precise, embodying cautions suggested by this writer in reviewing Dvorine's earlier work in this JOURNAL (58, 1945, 276-278). These cautions relate to constant distance, timing, illumination, angle of plate, and Dvorine's neglect of the critical blue-green and purplish-red hues. In the 1953 test, the Ishihara device of a duplex diagnostic plate for separating deuteranoid from protanoid, featuring on the right the purplish red which is colorless for the true deuteranoid, on the left the spectral red which darkens to gray in the extreme protanoid, is adopted, using, however, gray digits on color, rather than colored digits on gray, possibly an advantage, yielding more clear-cut results (depending on the accuracy of the inks used).

As for *scoring*, no statistics or table of responses for the different plates are offered. We are asked to accept the designer's dictum that incorrect responses to 3 of the 14 plates indicate red-green deficiency. One to two errors may be due merely to a mysterious "predilection" (one surmises that *antipathy* is intended) for certain *colors in combination*. It will be recalled that in 1944 Dvorine insisted that most of the individuals classified as 'color blind' can see all colors *singly* with ease. Only when hues are juxtaposed does a kind of paralyzing psychosis set in, inducing figure-ground confusion—a handicap which he optimistically offered to remove by training. That unsupported and quite unscientific notion is here apparently revived to explain the occasional illegibility of a plate to the otherwise normal. There is no recognition of the existence of variants from typical sensitivity (on a presumably histological or physiological basis) such as Pickford's work indicates. There is no finesse in selecting color-pairs to test it (as in the forthcoming 4-degree Hardy-Rand-Rittler Polychromatic Plates), no codifying of errors on certain plates, no attempt to correlate the findings with the scores on other tests. The simple unsupported claim is made that the test does not fail the normal or misclassify the defective. This reviewer, having made only limited laboratory observations with the plates, has at present no evidence to refute it. The colors paired, few of them highly saturated, appear genuinely pseudo-isochromatic for the average defective examined.

On the debit side, there are numerous shortcomings that disqualify the test for clinical or for exact scientific use. For example, there is looseness or carelessness in language and terminology. 'Similar' is used in reference to color-pairs where 'identical' is intended; 'dark tints and light tints,' where saturated and desaturated are the truer designations (there are actually no 'dark tints' in the plates). In the revolving nomenclature chart, with a 1-in. circular exposure-window showing in turn each of 8 saturated (or 'dark tint') colors on one side, and on the other, tints of the same, the hue designated as 'purple' (5*P*, it is true, on the unsatisfactory Munsell circle) is really a reddish violet, and is listed as *Vt* (or Violet) on the record-sheet for Plates 14, 15, and 23 (violet dots on a blue ground). In

the duplex plates diagnostic of protanoid vs. deuteranoid, 6, 7, and 19 (the latter wrongly printed 23), a reddish purple is designated as *Ma* (for Magenta?) without explanation. There is no stabilization of terms, no recognition of the fact that *violet* is a spectral hue, and *purple* an off-spectrum mixture of red and violet.

While there is no recurrence of the somewhat heterodox 1944 asseveration that there are three primary colors, red, blue, and yellow, three secondary colors obtained by mixture, with brown a mixture of all three primaries, and gray, their absence (ideas borrowed perhaps from the printer, not from science), there is little promise of the scientific usefulness of the test. No guarantee of the purity or constancy of the inks used is offered. One suggests that the sphere of the test is the high school course in physiology, or the optometry school, to illustrate the confusion-colors of the *average* defective—orange with olive-green, blue with violet, green with greenish yellow, light brown or buff with scarlet, purple with green or gray for the deuteranoid, red with dark gray for the protanoid—all readily explainable in terms of any good four-primary color system. With the Ishihara (or *AO*) plates, each of which commonly uses three to four hues, this simple demonstration of the 'metamers' or pseudo-isochromatic hues, a practical understanding of which by the normal-visioned is highly desirable, is impossible. Such a laboratory use should, of course, be under the direction of a competent instructor, able to caution against acceptance of the test-results as proof of normalcy or deficiency without further corroboration.

A further use of the test is in demonstrating the need in scientific work of controlled conditions in distance, illumination, timing, and so forth. Munsell papers should be substituted for the inked colors used in the nomenclature disk. The plates, mounted on too-thin cardboard, tend to stick together, and should be equipped with 1-in. gummed gray tabs to be manipulated more readily, and, since they are unnumbered, an identifying letter should be inscribed on the back of the tab. The loose-leaf rings would operate more easily if mounted on the lower, not the left-hand edge.

Cornell University ELSIE MURRAY

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXVIII SEPTEMBER, 1955 No. 3

STIMULUS-RESPONSE ASYNCHRONYSM AND DELAY OF REINFORCEMENT IN SELECTIVE LEARNING

By RICHARD A. DURYEA, University of Texas

The course of simple selective learning—the rate at which the tendency for a stimulus (S) to elicit some response (R) is changed by the motivational consequences of that response (G)—is markedly influenced by the temporal relations among the three critical events. The interval between S and R has been studied under the headings of *delayed response* and *stimulus-response asynchronism,*[1] while the interval between R and G has been studied in experiments on the *temporal gradient of reinforcement.*[2] Theoretical interest in the latter problem has recently given rise to the question of the relation between the two intervals—a question which the present experiment was designed to answer. It represents the first systematic comparison of the effects of short S–R and R–G intervals on the rate of simple selective learning.

The problem of the temporal gradient of reinforcement arose from the work of

* Accepted for publication November 1, 1954. This article is based on a doctoral dissertation presented to the University of Texas. The writer is indebted to Professor M. E. Bitterman who directed this research.

[1] J. T. Cowles, Delayed response tested by three methods and its relation to other learning situations, *J. Psychol.*, 9, 1940, 103-130; C. L. Hull, *Principles of Behavior,* 1943, 165-182; W. E. Kappauf and Harold Schlosberg, Conditioned responses in the white rat: III. Conditioning as a function of length of delay, *J. Genet. Psychol.*, 50, 1937, 27-45; W. G. McAllister, A further study of the delayed reaction in the albino rat, *Comp. Psychol. Monog.*, 8, 1932, 1-103; F. McCord, The delayed reaction and memory in rats: I. Length of delay, *J. Comp. & Physiol. Psychol.*, 27, 1939, 1-37; H. M. Wolfe, Conditioning as a function of the interval between the conditioned and the original stimulus, *J. Gen. Psychol.*, 7, 1932, 80-103.

[2] Hull, The goal gradient hypothesis and maze learning, *Psychol. Rev.*, 39, 1932, 25-43; *Principles of Behavior,* 1943, 135-164; C. T. Perin, A quantitative investigation of the delayed reinforcement gradient, *J. Exper. Psychol.*, 32, 1943, 37-51; J. B. Wolfe, The effect of delayed reward upon learning in the white rat, *J. Comp. Psychol.*, 17, 1934, 1-21.

Thorndike, who suggested that a temporal delay interposed between a response and its motivational consequences (effect) would retard learning.[3] Shortly thereafter, experiments by Watson with the problem-box,[4] and by Warden and Haas with the multiple T-maze,[5] failed to demonstrate a clear decrement in the rate of learning with increasing delay of reinforcement. In each of these studies *Ss* were delayed in the feeding compartment of the apparatus where food was present in a dish with a perforated cover. In subsequent experiments by Hamilton,[6] Roberts,[7] and others,[8] some evidence of an inverse relation between rate of learning and delay of reward was obtained, but not until Wolfe's study in 1934 did a consistent temporal gradient appear.[9] Wolfe described the gradient of reward as falling rapidly within the first minute and then gradually levelling off, with some learning taking place even at 10 and 20 min.

In none of the early studies was the problem of secondary reinforcement recognized. If neutral stimuli associated with reward thereby acquire rewarding properties, it becomes necessary in studies of the gradient of reward to eliminate, during the delay period, not only all primary rewards but all stimuli which may have been associated with them. Perin, working under Hull's guidance, made the first attempt to control the effects of secondarily reinforcing stimuli.[10] He employed a modified Skinner-box on the assumption that, within a single compartment serving as stimulus-situation, delay-chamber, and food-box, there could be no *differential* secondary reinforcement of the lever-depression response which was to be learned. The results obtained by this method contrasted markedly with those of previous studies, showing a sharp, negatively accelerated gradient, which, when extrapolated, fell to zero at a value near 30 sec. It was largely on the basis of this evidence that Hull reformulated his earlier goal-gradient hypothesis,[11] and distinguished between primary and secondary gradients of reinforcement. "What was originally regarded as a single principle," he wrote, "has turned out upon intensive investigation to involve two fairly distinct principles: (1) the short gradient reported by Perin, which will be called the *gradient* of reinforcement and (2) the more extended *goal gradient* which is presumably generated as a secondary phenomenon from Perin's gradient of reinforcement acting in conjunction with the principle of secondary reinforcement."[12]

---

[3] E. L. Thorndike, *Educational Psychology*, Vol. 1: *The Original Nature of Man*, 1913, 173.

[4] J. B. Watson, The effect of delayed feeding upon learning, *Psychobiol.*, 1, 1917, 51-60.

[5] C. J. Warden and E. L. Haas, The effect of short intervals of delay in feeding upon speed of maze learning, *J. Comp. Psychol.*, 7, 1927, 107-116.

[6] E. L. Hamilton, The effect of delayed incentive on the hunger drive in the white rat, *Genet. Psychol. Monog.*, 5, 1929, 131-207.

[7] W. H. Roberts, The effect of delayed feeding on white rats in a problem cage, *J. Genet. Psychol.*, 37, 1930, 35-58.

[8] A. C. Anderson, Time discrimination in white rats, *J. Comp. Psychol.*, 13, 1932, 27-55; C. F. Sams and E. C. Tolman, Time discrimination in white rats, *J. Comp. Psychol.*, 5, 1925, 255-263.

[9] Wolfe, *op. cit.*, 1-21.

[10] Perin, *op. cit.*, 37-51; The effect of delayed reinforcement upon the differentiation of bar responses in the white rat, *J. Exper. Psychol.*, 32, 1943, 95-109.

[11] Hull, The goal gradient hypothesis and maze learning, *loc. cit.*, 25-43.

[12] Hull, *Principles of Behavior*, 1943, 143.

According to Hull, the development of the goal-gradient, such as the one obtained by Wolfe,[13] involves the following series of events: first, the stimuli falling within the primary gradient of reinforcement from the point of primary reward acquire secondary reinforcing properties; these stimuli, in turn, having their own short gradients, lend secondary reinforcing properties to immediately antecedent stimuli; and so on. Thus, the goal-gradient is derived from the summation of numerous overlapping primary gradients. Hull explained the results of earlier studies by Watson, by Warden and Haas, and by Wolfe in terms of this concept.

In a subsequent analysis of the rôle of secondary reinforcement in delayed-reward learning, Spence carried Hull's analysis to its logical conclusion.[14] He rejected entirely the concept of a primary gradient of reinforcement and attempted to account for all learning in the absence of immediate primary reinforcement on the basis of immediate secondary reinforcement. Spence emphasized that interoceptive as well as exteroceptive stimuli may acquire secondary reinforcing properties, and he attributed the results obtained with the Skinner-box by Perin[15] and by Perkins[16] to their failure to eliminate interoceptive sources of secondary reinforcement. He argued that certain proprioceptive consequences of the lever-depressing response, for example, might reasonably be expectd to persist during the delay-interval and thus to acquire secondary reinforcing properties as a result of contiguity with the subsequent primary reinforcement. To the extent that the proprioceptive traces become secondarily reinforcing, the response to be learned is immediately reinforced. The gradient of reinforcement obtained in the Skinner-box could thus be attributed, not to a primary gradient of reinforcement, but to a gradient of stimulus-generalization. The greater the delay, the greater the difference between the afferent conditions immediately following the response and the afferent conditions immediately contiguous with primary reinforcement (on the assumption that the proprioceptive traces of the response change in time), and, consequently, the weaker the secondary reinforcing effect of the immediate afferent consequences of response. Under these conditions, the maximum delay of primary reward at which some learning will take place depends on the length of time during which the changing proprioceptive trace remains similar enough to the proprioceptive feedback to permit some generalization of reinforcing properties.

Grice tested this formulation in a simple black-white discrimination which was designed to eliminate differential secondary reinforcement arising from proprioceptive traces.[17] With the position of the positive cue randomly varied, right and left turns were equally often rewarded, and the proprioceptive consequences of the two responses could thus be expected to acquire equal secondary reinforcing properties. Under these conditions the development of a preference for the correct, as compared with the incorrect, turn in the presence of a given arrangement of the stimuli could not be attributed to differential proprioceptive secondary reinforcement. The

---

[13] Wolfe, *op. cit.*, 1-21.

[14] K. W. Spence, The role of secondary reinforcement in delayed reward learning, *Psychol. Rev.*, 54, 1947, 1-8.

[15] Perin, A quantitative investigation of the delayed reinforcement gradient, 37-51.

[16] C. C. Perkins, The relation of secondary reward to gradients of reinforcement, *J. Exper. Psychol.*, 37, 1947, 377-392.

[17] G. R. Grice, The relation of secondary reinforcement to delay reward in visual discrimination learning, *J. Exper. Psychol.*, 38, 1948, 1-17.

temporal gradient obtained by Grice was much steeper than that obtained by Perin, falling close to zero at about 10 sec. This brief gradient was attributed to the one "remaining possibility of secondary reinforcement in this type of learning problem." As Grice explained, "So long as a trace of the positive stimulus is effective at the time of reward, this trace may become a secondary reinforcing agent, and the positive stimulus itself might then come to provide immediate secondary reinforcement. The limit within which delayed reward learning could take place in a visual discrimination problem would depend then on the time that the after-effects (stimulus-traces) of the positive and negative stimuli remain discriminably different."[18]

To test this interpretation, Smith attempted to make independent measurements of the rate at which the traces of black and white stimuli decay in time.[19] Using a simple T-maze, Smith trained his animals to turn in one direction when the approach to the choice-point was white and in the opposite direction when the approach was black, interposing delays of varying duration between the termination of the stimulus (the time of departure from the critically colored area) and the turning response. Assuming that learning would be possible under these conditions only as long as the trace of the relevant stimulus remained discriminable at the time of choice, and assuming the correctness of Grice's interpretation, Smith deduced that his own experiment (with S–R delay) would yield a temporal gradient like that obtained by Grice (with R–G delay). Although Smith appeared to be satisfied with the similarity of the two gradients—a similarity achieved by rather questionable manipulation of the ordinates—both his experiment and his theoretical analysis leave much to be desired.

It should be obvious, first of all, that Smith's experimental conditions were not comparable with those of Grice. Grice's animals learned a simple simultaneous discrimination, which required only that the animals approach that one of the two stimuli which was consistently reinforced. Smith used a successive discrimination in which the two components were not differentially reinforced, and in which learning could take place (according to Spence's theory) only on the basis of some sort of conditionality of brightness and position. While it is true that the simultaneous situation does not permit the control of the *S–R* delay which is possible in the successive situation, the successive situation permits effective control *both* of S–R and R–G delays. It seems obvious, therefore, that Smith should have compared the two delays in the successive situation, instead of comparing S–R delay in the successive situation with R–G delay in the simultaneous situation. Furthermore, the *Ss* employed by Smith were very different from those employed by Grice. Grice's animals were albino rats, while the majority of Smith's were hooded, a difference which cannot be tolerated in the face of Smith's own finding that the performance of the hooded animals was quite superior to that of the albinos.

The experiment to be reported here provides a more adequate comparison of the relative effects on learning of S–R and R–G delays. A successive black-white discrimination was employed to permit the intro-

[18] *Ibid.*, 3.

[19] M. P. Smith, The stimulus trace gradient in visual discrimination learning, *J. Comp. & Physiol. Psychol.*, 44, 1951, 154-160.

duction of both kinds of delay without changing the fundamental nature of the learning problem. The apparatus devised for this purpose was a modified T-maze which contained delay-boxes between the choice-point and each goal-compartment (R–G delay-boxes) as well as between the stimulus-compartment and the choice-point (S–R delay-box). The animals were rewarded for turning *right* when the stimulus-compartment was *black* and for turning *left* when the stimulus-compartment was *white*. Delays of 2 or 4 sec. were imposed either before (S–R) or after (R–G) the choice-point.

Spence and Grice have not dealt specifically with delay of reinforcement in successive discrimination, which poses some special problems. One important feature of the successive situation is that no afferent component, either visual or proprioceptive, acquires differential reinforcing properties, since both turns and both visual cues (or their traces) are equally often associated with primary reinforcement. If a temporal gradient occurs in the successive problem (as Smith's data suggest), and if this gradient is to be interpreted in the same way as that obtained by Grice, differential secondary reinforcement must be sought in visual-proprioceptive *compounds*. While the trace of white alone and the trace of black alone do not differ in secondary reinforcing properties, and while the same is true of the proprioceptive traces of right and left turns, the compound traces of white-left and black-right should acquire greater secondary reinforcing properties than the traces of white-right and black-left. This analysis suggests (contrary to Smith's interpretation) that the functional consequences of S–R and R–G delays should be quite different.

In Grice's study, which employed a simultaneous discrimination, the positive stimulus (white) presumably acquired secondary reinforcing properties which in turn provided immediate reinforcement to the appropriate response and thus permitted learning under the conditions of R–G delay. The longer the delay between the occurrence of the positive stimulus and the primary reinforcement in this situation, the weaker should be the secondary reinforcing properties of the stimulus (due to generalization), and, consequently, the less rapid the learning. In the same manner, learning in the successive problem should depend on the temporal separation of the critical *compounds* (alleged to provide immediate secondary reinforcement) from the source of primary reinforcement. An extension of the Spence-Grice analysis to the successive situation suggests, therefore, that immediate secondary reinforcement of a correct response should be greater for an S–R delay than for an R–G delay of

equal duration, since in the former case there is greater temporal contiguity between the afferent state of affairs which immediately follows (and secondarily reinforces) the response and the afferent state of affairs which exists at the moment of primary reinforcement. In the R–G situation the delays are imposed *after* the response, whereas in the S–R situation the response is *immediately* followed by primary reinforcement.

It cannot, however, be concluded from this analysis that learning should be more rapid with S–R delay than with R–G delay of equal duration. While the stimuli to be discriminated may have the secondary reinforcing function which the Spence-Grice interpretation suggests, they also have a response-*eliciting* function which must be taken into account. In a successive discrimination, response at the choice-point is guided by an afferent pattern of which the trace of the stimulus-compartment is an essential component. Since the traces of the discriminanda must be assumed to fade in time, and since weak traces of black and white are less discriminable than strong traces of black and white, the S–R delay presents a more difficult problem than the R–G delay. From the point of view of secondary reinforcement, therefore, learning should be more rapid for S–R delay, while from the point of view of response-elicitation, learning should be more rapid for R–G delay. It follows, then, that the results obtained in any systematic comparison of the two types of delay would depend on the relative strength of the two effects and on the shapes of the two temporal functions. It is conceivable that the two effects might exactly cancel each other over the entire range of measurable time-intervals, but (contrary to Smith) the Spence-Grice interpretation does not necessarily imply that rate of learning must be equally affected by S–R and R–G delays of equal duration.

It is interesting to note that from a simple cognitive point of view, which minimizes the functional rôle of response, S–R and R–G delays of equal length might be expected to have equivalent effects on the course of successive discrimination. Assuming that solution of the problem requires the animal to learn a relation which involves three aspects of the experimental situation—between the occurrence of a particular *stimulus* and the presentation of a *reward* in a particular *place*—solution (the integration of the three aspects) should be retarded by a given delay irrespective of whether it is interposed before or after the intervening response.

A second feature of the present study was designed to pursue further the analogy between Grice's simultaneous problem and the successive problem here employed. In Grice's experiment the effect of immediate

differential secondary reinforcement was examined by the use of appropriate black and white goal-boxes. The animals in this phase of his study either approached white, were delayed, and then entered a white goal-box containing food or they approached black, were delayed, and then entered a black goal-box containing no food. Grice predicted that the whiteness of the goal-box (present at the moment of primary reinforcement) should acquire strong secondary reinforcing properties, which in turn should generalize to the whiteness of the positive stimulus and thus facilitate learning (relative to a control condition in which both goal-boxes were gray). This prediction was confirmed by the results. To obtain a rough approximation of the generality of this interpretation, a somewhat analogous goal-box manipulation was used in the context of the present successive discrimination. The apparatus used was the same as that of the major portion of the study, except that one of the goal-boxes was white and the other was black. Two groups of animals were trained with a 4-sec. R–G delay to go right to black and left to white. For one of the groups (Congruent), the right goal-box was always black and the left goal-box was always white. Thus, the compounds allegedly providing the immediate secondary reinforcement thought to be essential for learning (black-right and white-left) were in this group directly associated with primary reinforcement. For the second group (Incongruent), the right goal-box was always white and the left was always black. In this case, the compounds opposite to those allegedly responsible for the immediate secondary reinforcement of correct responses (black-left and white-right) were directly associated with reinforcement. The Spence-Grice interpretation leads to the prediction that the Congruent condition should produce more rapid learning than the Incongruent condition. Precisely the same outcome is expected on the simple cognitive assumption that what the animal learns in this problem is the relation between the occurrence of a particular stimulus and the presentation of reward in a particular locality.

Method

Subjects. The *Ss* of this experiment were 60 naïve albino rats, 30 males and 30 females, ranging in age from 90 to 120 days. They were divided into six groups of 10 animals each which were designated, respectively, SR_2, SR_4, RG_2, RG_4, CRG_4, and IRG_4.

Apparatus. The apparatus employed is illustrated in Fig. 1. It was a simple T-maze constructed of seven units, each of which was 12 in. long, 5 in. high, and 4 in. wide. The stem of the maze consisted of three units—a starting chamber, a stimulus-chamber, and an S–R delay-chamber terminating at the choice-point. Each arm of the maze consisted of two units, an R–G delay-chamber and a goal-chamber. The

starting chamber (S) and the three delay-chambers were identical units, painted flat gray. The stimulus-unit and the two goal-units were variable in color. Each of these units consisted of a set of three chambers (black, gray, and white), any one of which could be made an integral part of the maze-path on any given trial. The goal-units were interchangeable, one being used for reinforcement and the other for nonreinforcement throughout the study. Each chamber of the stimulus-unit was equipped with a separate guillotine-door (D_2) painted the same color as the rest of the chamber. Other guillotine-doors, designed either to delay the animal's progress through the maze or to prevent retracting, are shown in Fig. 1 (D_1, D_3, D_4, D_5, D_6, and D_7). These were painted flat gray. All doors were operated by *E* from a position just back of the starting box with the aid of a string-and-pulley system. Fig. 1 also shows gray curtains (C) at the entrance to each goal-unit

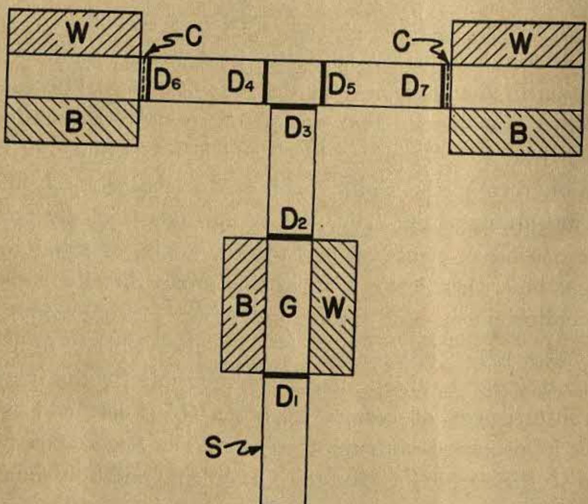


FIG. 1. DIAGRAM OF THE APPARATUS
W, white; B, black; G, gray; C, curtain; D, door; S, starting compartment.

which concealed the contents of the goal-chamber prior to the entry of the animal. The entire maze was covered with ½ in. wire mesh. The apparatus was set on a sturdy table in the homogeneous environment provided by the Dome Room of the animal laboratory.

Preliminary training. All *Ss* were first isolated in individual cages which contained unlimited supplies of food and water. At the end of two weeks they were put on a restricted diet which gradually reduced each *S* to 80% of its satiated bodily weight. This level was maintained throughout the experiment, each animal being given enough food at the end of each daily series of trials to bring it to 80% of its satiated weight at the beginning of the next day's runs.

There were four days of preliminary training. On the first day *S* was allowed to explore the wings of the maze for several minutes. Both goal-chambers contained

small amounts of wet mash. On the second and third days *S* was given 10 complete runs through the maze, from the starting box to one of the goal-boxes. On odd-numbered trials it was free to make either a right or left response, but on even-numbered trials it was forced (by means of D_4 or D_5) to go to the side opposite that chosen on the previous free trial. This procedure provided equal experience with both arms of the maze. The fourth day of training was identical with the second and third except that *S* was given experience with the guillotine-doors to be used in the experiment proper. Doors D_1, D_2, D_3, D_6, and D_7 were down when the animal was placed in the starting box and then raised in quick succession as the animal progressed through the maze. Doors D_4 and D_5 were used to prevent retracing.

The gray section of the stimulus-chamber was used throughout preliminary training. The goal-boxes were also gray for groups SR_2, SR_4, RG_2, and RG_4, as they were during the experimental training. Group CRG_4, in this and in subsequent stages of the experiment, found a black goal-box on the right and a white goal-box on the left. For group IRG_4, the right goal-box was white and the left one was black.

Experimental training. All *Ss* were trained to go right when the stimulus-chamber was black and left when it was white. The order of stimulus-presentation, BWBBWWBWWB/WBWWBBWBBW, was repeated every 20 trials. On each trial *S* was confined in the stimulus-box for 5 sec. to insure effective stimulation.[20] Ten free-choice, non-correction trials a day were given to all animals with the exception of some in the Congruent and Incongruent groups (to be described later). The animal time between consecutive trials for each *S* was 7 min. Throughout the training a small amount of wet Purina mash was used for incentive, and the possibility of differential olfactory cues was controlled by placing an equal amount of food just outside the non-reinforced goal-box. This procedure was continued until a criterion of 18 out of 20 correct choices was attained, although it was necessary to terminate the training of some *Ss* before the criterion was reached. In the actual conduct of the experiment, limitations of time made it possible to run only 30 of the 60 animals at once. Thus, 30 animals 5 in each of the 6 experimental groups, were run in the first 40 days of the project and the remainder in the next 40 days.

Four of the six experimental groups (SR_2, SR_4, RG_2, and RG_4) were used for a systematic comparison of the effects on learning of S–R and R–G delays. On each trial the animals in groups SR_2 and SR_4 were delayed for 2 sec. and 4 sec. (respectively) in the compartment leading to the choice-point. As the animal left the stimulus-compartment, D_2 was closed, and 2 or 4 sec. later D_3 was raised. After the right or left turning response, D_4 or D_5 was used to prevent retracing (depending on the direction of the response). For these S–R delay animals, D_6 and D_7 were always open, although gray curtains concealed the contents of the goal-boxes. It will be noted that the apparatus made it possible to control only the minimal delay—although the turning response could not be made before the 2-sec. or 4-sec. delay

---

[20] Abram Amsel, Rate of learning a visual brightness discrimination as a function of discriminanda durations, *J. Comp. & Physiol. Psychol.*, 45, 1952, 341-350; K. F. Muenzinger and F. M. Fletcher, The effect of an enforced delay at the point of choice in the visual discrimination habit, *J. Comp. Psychol.*, 23, 1937, 383-392.

had elapsed, there was no way of insuring that the animal would not delay *longer* than this imposed minimum time. During the preliminary training, however, the highly motivated animals learned to run rapidly through the maze, and the minimal delays subsequently imposed by the experimental procedure closely approximated the actual delays. The animals would claw at the delay-door until it was raised.

Groups RG_2 and RG_4 were delayed in the compartments between choice-point and goal-box for 2 and 4 sec. respectively. For these animals, D_3 was always open but D_6 and D_7 were closed and remained so until the required interval had elapsed. D_4 and D_5 were used to prevent retracing. Delays were imposed both after correct and after incorrect responses. All delay-times were measured from the moment the animal entered the delay-chamber.

Groups CRG_4 and IRG_4 were treated in all respects like group RG_4 with the following exceptions: For group CRG_4 (the Congruent group), the right goal-box was black and the left goal-box was white on all trials. For group IRG_4 (the Incongruent group) the right goal-box was white and the left goal-box was black. In the first half of the experiment, in which 5 animals of each group were run, a marked difference appeared between the Congruent and Incongruent groups. In an attempt to analyze this difference, the conditions for the first subgroups (designated $CRG_{4\text{-free}}$ and $IRG_{4\text{-free}}$) were changed for the second set of 5 animals in each of the two groups (subgroups $CRG_{4\text{-forced}}$ and $IRG_{4\text{-forced}}$). During the first 10 days of their training, only 5 of the 10 trials in each daily series were free. The rest were forced in such a way as to equate experience with both arms of the maze. On day 11 and all following days, all trials were free.

Results and Discussion

The data for all animals in all experimental conditions are shown in Table I in terms of total errors and total trials to reach the criterion (or, for those *Ss* which did not learn, total errors in 400 trials). Since the distributions obtained were not consistently normal, Wilcoxon's non-parametric method for unpaired replicates was employed in the analysis of differences between groups.[21]

S–R versus R–G delay. In Fig. 2 the course of learning in the four simple delay-groups is plotted in terms of mean errors per block of 20 trials. As each *S* reached criterion its training was terminated, but for the purposes of plotting these curves the assumption was made that it continued to perform errorlessly on succeeding days. Comparable rates of learning may be seen for group SR_2, RG_2, and RG_4, while group SR_4 learns more slowly than all three. In Table I the results for the four groups are summarized in terms of total errors and total trials to the criterion. By Wilcoxon's method, no significant differences appeared among the three best groups, but each of the three was significantly

[21] Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Co., 1949, 1-16.

superior to group SR_4 in terms of both measures at a level of confidence beyond 1%. These results, obtained by imposing both delays in a common learning situation, clearly suggest, therefore, that S–R delays tend to have a more deleterious effect upon learning than do corresponding R–G delays.

These findings fail to provide support for a simple cognitive theory. The difference in the effects of corresponding S–R and R–G delays suggests that the solution of the problem is not based merely on an integration

TABLE I

NUMBER OF TRIALS AND ERRORS REQUIRED BY EVERY ANIMAL IN EACH GROUP TO REACH THE CRITERION

	SR_2		SR_4		RG_2		RG_4	
	trials	errors	trials	errors	trials	errors	trials	errors
	70	24	120	49	70	21	90	33
	110	45	170	64	80	25	100	37
	150	61	280	111	80	32	110	40
	160	53	310	143	100	39	150	64
	160	54	380	159	140	58	170	77
	160	62	390	168	180	68	180	57
	230	106	400+	166+	210	87	180	71
	260	93	400+	172+	220	89	230	93
	270	110	400+	172+	270	115	290	79
	310	120	400+	197+	280	103	370	139
Mdn.	160	61.5	385	163.5	160	63.0	175	67.5
Mean	188	72.8	325+	140.1+	163	63.7	187	69.0
	CRG_{4-free}		IRG_{4-free}		$CRG_{4-forced}$*		$IRG_{4-forced}$*	
	trials	errors	trials	errors	trials	errors	trials	errors
	110	37	90	30	140	52	80	28
	130	55	90	34	210	61	80	28
	220	72	110	42	210	77	130	50
	220	95	120	45	230	76	160	50
	280	95	140	53	330	101	180	53
Mdn.	220	72.0	110	42.0	210	76.0	130	50.0
Mean	192	70.8	110	40.8	224	73.4	126	41.8

* The data shown for the forced training conditions are based on the free trials only.

of afferent events separated in time—the occurrence of a given stimulus and the presence or absence of food in a given end-box. Contrary to the cognitive view, the response which mediates the sequence of experiences in this situation must enter into the learning in an integral fashion since its temporal position (early in the case of R–G delay or late in the case of S–R delay) markedly influences the course of discrimination. Nor is the validity of the Spence-Grice position, which accords a central rôle to response (proprioceptive traces compounded with traces of the discrimi-

nanda provide the immediate reinforcing conditions said to be essential for learning), directly apparent in the present data. Retardation in the SR_4 group may be attributed to the rapid fading in time of the traces of the discriminanda relative to change in the secondarily reinforcing compounds, but this explanation is not wholly satisfactory. How, it might be asked, do the secondarily reinforcing compounds remain sufficiently discrete to promote rapid learning in group RG_4 when the traces of the critical black and white components do not remain discrete enough to promote rapid learning in SR_4? The results of the second phase of the experiment provide further evidence of the inadequacy of both interpretations.

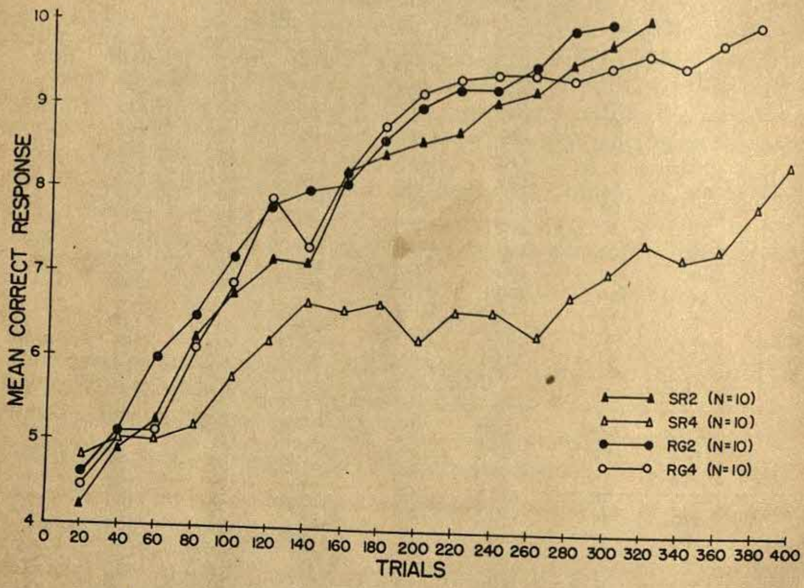


FIG. 2. LEARNING CURVES FOR THE FOUR DELAY GROUPS

Congruent vs. Incongruent groups. The course of learning in the Congruent and Incongruent groups is shown in Fig. 3. As in Fig. 2, the assumption was made in plotting the curves that the animals continued to perform perfectly after reaching criterion, although no further trials were given. Fig. 3 shows mean correct responses in each block of 20 trials, except that each of the first 5 points of each curve for the subgroups given forced training is *computed* on the basis of the 10 free trials in the block of 20. Superior rates of learning are seen for both free and forced Incongruent groups relative to the corresponding Congruent groups, Table

I summarized these results in terms of total errors and trials to criterion, *based on free trials only*. By Wilcoxon's test, all Incongruent *Ss* combined were significantly superior, in terms of both measures, to all Congruent *S*s at a level of significance beyond 1%. By the same test, each of the Incongruent subgroups was superior to the corresponding Congruent subgroup at the 5-% level.

This outcome seems impossible of explanation in terms of the Spence-Grice position. If the successive solution depends on the extent to which the compounds black-right and white-left acquire reinforcing properties

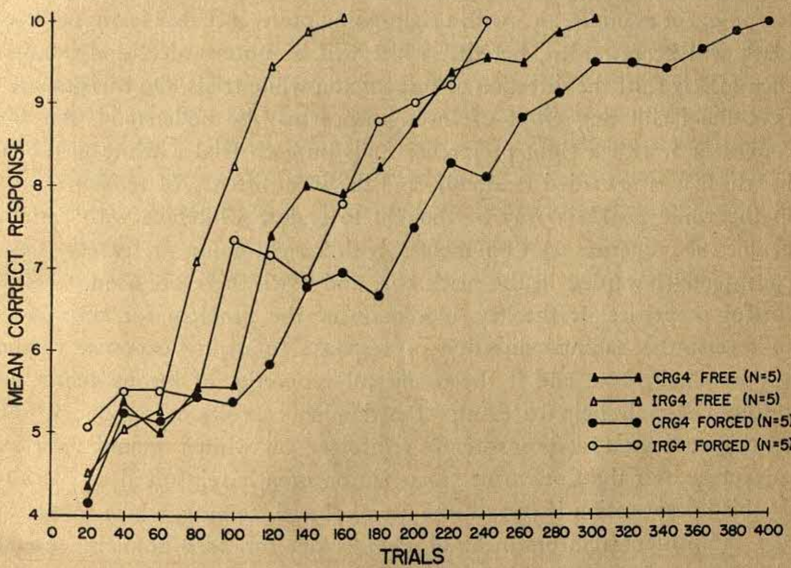


FIG. 3. LEARNING CURVES FOR GROUPS CRG_4 AND IRG_4 UNDER FREE AND FORCED CONDITIONS

relative to the compounds white-right and black-left, the Congruent groups (in which black-right and white-left were directly associated with primary reinforcement) should learn more rapidly than the Incongruent groups (in which black-left and white-right were directly associated with primary reinforcement). The results, however, are diametrically opposed to this expectation. Such data are obviously contrary also to expectations derived from a cognitive analysis based on structural considerations.

The experimental design, it should be noted, did not initially involve the use of forced training; all choices made by the first five animals in the

Congruent and Incongruent groups were free. In attempting to understand the clear superiority of the Incongruent condition which appeared in this first phase of the work, however, an hypothesis was made which could be tested by the introduction of some forced trials in the training of the remaining five *S*s in each group. Since on the early trials the *S*s ran almost exclusively to a preferred position, and since mastery of the problem seemed to begin with a shift in the direction of response on trials in which response in the preferred direction would have been unreinforced, the hypothesis was entertained that the events of *unrewarded* trials during the dominance of positional preferences were of central importance for learning. For example, an *S* with a right-going preference may learn, not that black will be rewarded, but that white will be unrewarded, and it must then merely shift the direction of response on white trials. On this assumption, the facilitating effect of incongruence may be understood. An Incongruent *S* with a right preference finds on each trial a white goal-box in which it is rewarded randomly, and the inconsistency of reinforcement in the white goal-box may be thought to endow whiteness with certain conflictful properties. A Congruent *S* with a right-going preference is inconsistently rewarded in the black box, and black therefore acquires conflictful properties. If the key to solution of the problem for both *S*s is to reverse the habitual direction of response following exposure to the white stimulus-box, and if the conflictful properties of inconsistently reinforced stimuli help to disrupt the dominant positional tendency, the Incongruent group (inconsistently reinforced on white) should have an advantage over the Congruent group (inconsistently reinforced on black). If this interpretation is correct, the use of forced training, which results in the inconsistent reinforcement of *both* colors for *both* animals, should deprive the Incongruent condition of systematic advantage. The results for the forced subgroups make it necessary, however, to reject this interpretation. While the forced training had a general retarding effect on the learning of both groups, the superiority of the Incongruent condition was maintained. Further systematic analysis of the color relations is necessary before an adequate account of the present results is possible.

SUMMARY

The principal purpose of the present study was to compare the effects of stimulus-response (S–R) and response-reward (R–G) delays on the rate of selective learning. Four groups of rats were trained on a successive brightness discrimination in a modified T-maze. All animals were rewarded for turning right when the stimulus-compartment in the stem

was black and for turning left when the stimulus-compartment was white. Two of the groups were delayed immediately before the choice-point response (S–R delay), one for 2 sec. and one for 4 sec. The other two groups of animals were delayed in the arms of the maze immediately following the response (R–G delay) for the same periods of time. The results showed no difference between the S–R and R–G groups delayed for 2 sec., but with a delay of 4 sec. the R–G group was significantly superior to the S–R group. Supplementary data on the effect of manipulating goal-box colors also were obtained. Animals rewarded in a white goal-box on the left for turning left to a white stimulus, and in a black goal-box on the right for turning right to a black stimulus (Congruent condition), were significantly inferior in rate of learning to animals rewarded in a black-goal box on the left for turning left to a white stimulus and in a white goal-box on the right for turning right to a black stimulus (Incongruent condition). These conditions were directly analogous for the successive discrimination to conditions employed in previous work with simultaneous discriminations. The results obtained suggest the lack of generality of the theoretical position developed to account for the earlier results on the delay of reinforcement.

RATE OF APPARENT CHANGE IN A DYNAMIC AMBIGUOUS FIGURE AS A FUNCTION OF OBSERVATION-TIME

By KENNETH T. BROWN, Wright Air Development Center, Ohio

The apparent changes which occur with an ambiguous figure, such as the Necker cube, have been studied in relation to stimulus-variables.[1] The rate of apparent change has also been tried as a measure of such diverse factors as visual fatigue, personality-traits, brain-damage, anoxia, and drug-effects.[2] In the present study, an improved method for measuring

* Accepted for publication July 14, 1954. From the Aero Medical Laboratory. The author is indebted to Virginia W. Brown for assistance in the collection and analysis of data.

[1] Kate Gordon, Meaning in memory and in attention, *Psychol. Rev.*, 10, 1903, 267-283; C. H. Graham, Area, color and brightness difference in a reversible configuration, *J. Gen. Psychol.*, 2, 1929, 470-481; C. O. Weber, Apparent movement in Lissajou figures, this JOURNAL, 42, 1930, 647-653; M. F. Washburn, H. Mallay, and A. Naylor, The influence of the size of an outline cube on the fluctuations of its perspective, this JOURNAL, 43, 1931, 484-489; W. T. Donahue and C. H. Griffitts, The influence of complexity on the fluctuations of the illusions of reversible perspective, this JOURNAL, 43, 1931, 613-617; M. F. Washburn and Annette Gillette, Motor factors in voluntary control of cube perspective fluctuations and retinal rivalry fluctuations, this JOURNAL, 45, 1933, 315-319; H. Goldhamer, The influence of area, position, and brightness in the visual perception of a reversible configuration, this JOURNAL, 46, 1934, 189-206; M. F. Washburn, C. Reagan, and E. Thurston, The comparative controllability of the fluctuations of simple and complex ambiguous perspective figures, this JOURNAL, 46, 1934, 636-638; B. R. Philip and V. R. Fisichelli, Effect of speed of rotation and complexity of pattern on the reversals of apparent movement in Lissajou figures, this JOURNAL, 58, 1945, 530-539; V. R. Fisichelli, Effect of rotational axis and dimensional variations on the reversals of apparent movement in Lissajous figures, this JOURNAL, 59, 1946, 669-675.

[2] I. E. Ash, Fatigue and its effects upon control, *Arch. Psychol.*, 4, 1922 (No. 31), 1-61; N. O. Frederiksen and J. P. Guilford, Personality traits and fluctuations of the outline cube, this JOURNAL, 46, 1934, 470-474; R. W. George, The significance of the fluctuations experienced in observing ambiguous figures and in binocular rivalry, *J. Gen. Psychol.*, 15, 1936, 39-61; W. F. Grether, J. T. Cowles, and R. E. Jones, The effect of anoxia on visual illusions, *AAF Aviat. Psychol. Program Res. Reports*, Report No. 19, 1947, 249-255; J. P. Guilford and K. W. Braly, An experimental test of McDougall's theory of extroversion-introversion, *J. Abnorm. & Soc. Psychol.*, 25, 1930, 382-389; H. L. Hollingworth, Perceptual fluctuation as a fatigue index, *J. Exper. Psychol.*, 24, 1939, 511-519; J. McV. Hunt and J. P. Guilford, Fluctuation of an ambiguous figure in dementia praecox and in manic depressive patients, *J. Abnorm. & Soc. Psychol.*, 27, 1933, 443-452; William McDougall, *Outline of Abnormal Psychology*, 1926, 441-449; May Smith, A contribution to the study of fatigue, *Brit. J. Psychol.*, 8, 1916, 327-350; M. F. Washburn, K. Keeler, K. B. New, and F. M. Parshall, Experiments on the relation of reaction-time, cube fluctuations, and mirror drawing to temperamental differences, this JOURNAL, 41, 1929, 112-117; G. K. Yacorzynski and Loyal Davis, An experimental study of the functions of the frontal lobes in man, *Psychosom. Med.*, 7, 1945, 97-107.

rate of apparent change has been developed, and further information has been obtained concerning the physiological process which causes the apparent changes.

It should be noted at the outset that all ambiguous figures seem to give rise to more than two different perceptions. Some figures give only two perceptions when they are viewed for short periods, but with prolonged observation other perceptions occur, at least with some *S*s. In the present experiment, *S* was required to press a key whenever the perception changed, regardless of whether the apparent direction of rotation reversed or whether there was some other apparent change. The measure used in this study may, therefore, be called *rate of apparent change* (*RAC*). Since this term is applicable to any ambiguous figure, regardless of the number or kinds of changes which occur, it will also be used in a general sense in this paper. The term is believed more useful than 'rate of reversal' or 'fluctuation rate' in dealing with ambiguous figures, since it has a broader meaning.

Theoretical background. Little theorizing has been done with regard to the physiological process underlying apparent changes in an ambiguous figure. Köhler and Wallach have advanced the hypothesis that the same physiological process causes both figural after-effects and apparent changes in an ambiguous figure, and they have called this physiological process "satiation."[3] Satiation may be defined broadly as a neurophysiological process which results from stimulation, and which tends to block any pattern of neural activity in the affected area of the visual system. The existence of such a process is indicated particularly by figural after-effects. Köhler and Wallach assert that when an inspection-figure is presented to one eye, the figural after-effect can be demonstrated in the other eye.[4] This finding means that the physiological process which causes figural after-effects is common to the two eyes, and Köhler and Wallach interpret it to mean that satiation occurs primarily in the visual cortex.

The hypothesis that both figural after-effects and apparent changes in an ambiguous figure are due to satiation is not dependent in any way upon the precise nature of the physiological process which constitutes satiation. Hence the term 'satiation' will not be used here to refer to the theory of Köhler and Wallach, as opposed to that of Osgood and Heyer,[5] regarding the exact nature of the physiological process of satiation. The term will be used only as defined above.

According to Köhler and Wallach, when an ambiguous figure like the Necker cube is seen in one perspective, satiation develops in the neurophysiological counterpart of that perception. Satiation tends to block the physiological process

---

[3] Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 270.

[4] *Ibid*, 271.

[5] C. E. Osgood and A. W. Heyer, Jr., A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118.

responsible for the perception, and when satiation reaches some critical level the perception changes and the other kind of perspective is seen. While satiation is developing for the second perception, it is presumably decaying for the first one. When some critical level of satiation is reached for the second perception, the first perception returns.

Since satiation is a process which results from stimulation, and which causes a decreased responsiveness to stimulation, it probably enters into many visual functions. Thus it is important to develop methods of measuring this process. According to Köhler and Wallach, both the figural after-effect and the rate of apparent change of an ambiguous figure should be measures of satiation. In view of the difficulties which have been reported in obtaining reliable measurements of figural after-effects, it seems reasonable that rate of apparent change may provide a better method of measuring satiation. Thus the present experiment was designed to provide some tests of the hypothesis that the same physiological process is responsible for both figural after-effects and apparent changes in an ambiguous figure.

The specific questions which this experiment was conducted to answer are the following: (1) What is the general shape of the curve showing *RAC* as a function of observation-time? (2) What is the reliability of such curves obtained by the present method? (3) What is the relation between curves for the left and right eyes of a given *S*? (4) How do the curves differ among *Ss*?

METHOD

Choice of ambiguous figure. The first problem in the study of ambiguous figures is what kind of figure to use. The static figures, such as the Necker cube, have disadvantages. The exact moment of apparent change is not always clear and there are periods during which the perception is hard to describe or classify. The latter difficulty is due partly to the fact that sometimes one part of the figure has one appearance, while another part has another appearance. It is also due partly to the occurrence of unclear perceptions. These difficulties become more pronounced with prolonged observation, and hence are particularly troublesome when *RAC* is being measured as a function of observation-time. Still another disadvantage of such figures is that the apparent changes which occur seem to be under a high degree of voluntary control.

Philip and Fisichelli overcame these disadvantages to a large extent by using Lissajous figures, which are two-dimensional sine-wave patterns produced on the face of a cathode-ray tube.[6] One part of such a pattern constantly moves to the right, while the other part constantly moves to the left, but the figure appears as a three-dimensional sine-wave pattern rotating around a vertical axis. When the sine-wave moving to the right appears in front of the one moving to the left, the pattern appears to rotate in one direction, but when the sine-wave moving to the right appears to be behind the other one, the apparent direction of rotation is reversed. Since the reversal of apparent depth is always accompanied by a reversal in the

[6] Philip and Fisichelli, *op. cit.*, 530-539.

apparent direction of movement, the reversal is always quite clear. Sometimes the perception changes to an entirely different kind of apparent movement, but the moment of change is well defined. Furthermore, all types of apparent movement form clear, unitary perceptions, which are relatively easy to describe and to classify. Another advantage of Lissajous figures is that there seems to be less voluntary control over the apparent changes than is the case with static figures.

Unfortunately, however, Lissajous figures also have certain disadvantages, the most important of which is that the physical rate of movement of the sine-waves is virtually impossible to hold constant. Since Philip and Fisichelli found that *RAC*

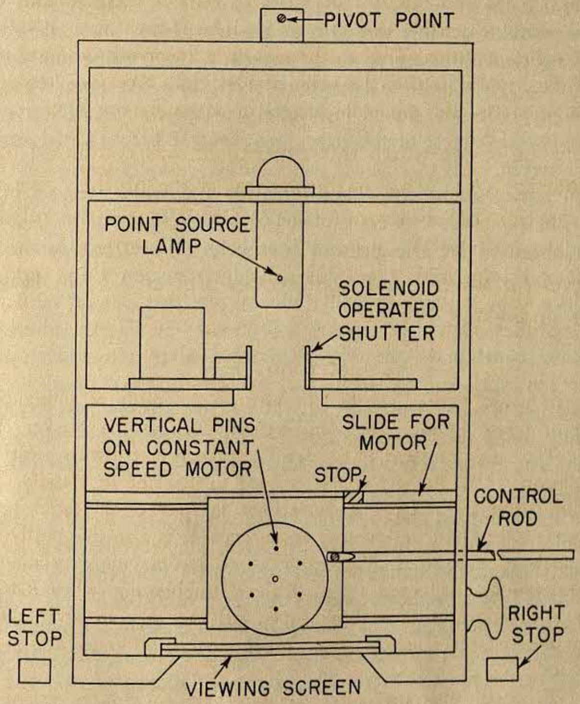


FIG. 1. SCHEMATIC DRAWING OF THE APPARATUS FROM ABOVE

varies directly with rate of movement, in measuring RAC as a function of observation-time it is necessary to hold rate of movement constant. A constant rate of movement was accomplished in the present study by developing a new type of ambiguous figure.

Apparatus. The device for presenting the ambiguous figure is shown in Fig. 1. The light source was a 25-w. concentrated arc lamp, which was essentially a point-source, since the light was emitted from a bead of metal 0.03 in. in diameter. The lamp and its associated power-supply were manufactured by the Western Union

Telegraph Company. In front of the lamp was an iris-type shutter, operated by a solenoid and controlled by a switch near *E*. The ambiguous figure was produced by six cylindrical brass pins, mounted vertically on a round metal plate and arranged in a circular pattern at 60° intervals. This plate was mounted on the shaft of a 15-r.p.m. constant-speed motor. The motor itself was mounted on a slide, so the pin-pattern could either be projected upon the screen at the front of the box, or moved to one side of the projection-path. A stop was provided for exact positioning of the pin-pattern when it was projected upon the screen, and *E* moved the entire unit by means of a rod which passed through the side of the box. Since a point-source of light was used, sharp shadows of the pins were projected upon the viewing screen at the front of the box. The screen itself had to provide a sharp image of these shadows, when viewed from the front by *S*, but at the same time it had to have sufficient diffusing power to prevent *S* from seeing into the apparatus and to provide a fairly uniform background luminance. An excellent combination of these two properties was found in Keuffel and Esser's 198 MX crystalline prepared tracing paper. A piece of this paper was mounted between two pieces of glass and used as a screen.

The screen was masked to provide a luminous area which was 3.6 in. wide and 2 in. high. The pin-shadows were centered along the horizontal axis of this area, and the shadows extended into the area from below. Since the shadow-pattern was 1.6 in. wide and 1 in. high, there was a luminous region 1 in. wide above and on both sides of the pattern. A small India-ink dot was put on the screen in the center of the shadow-pattern to serve as a fixation-point. Careful adjustments made all pin-shadows equal in height. The width of a given pin-shadow varied somewhat as the pin rotated, since rotation caused the distance from the pin to the point-source lamp to vary. The maximal width of the shadow was 8/64 in. and its minimal width, 6/64 in. Hence the time-averaged width of any given pin-shadow was 7/64 in., and the average width of all the shadows at any given time was also 7/64 in.

The lamp was stabilized at 115 v. The background luminance of the screen proved relatively uniform, and the luminance decreased symmetrically on all sides of the fixation-point. Background luminance was measured at the center of the pin-pattern with a Luckiesh-Taylor brightness meter, and the obtained value of 36 ft.-L. proved stable during the experiment. The shadow of a pin close to the screen was slightly darker than that of a pin farther away, probably due to less reflected light reaching the screen when the pin was close. The average brightness of near and far pins at the center of the pattern was 0.88 ft.L. An Eastman Color Temperature Meter was used to measure the color temperature of the screen, and a value of 3250°K. was obtained. The screen appeared definitely white.

This apparatus gave an illusion essentially like that which occurs with a Lissajous figure. The physical stimulus was two-dimensional, with one set of pin-shadows constantly moving to the right and another set constantly moving to the left, but it appeared as a three-dimensional set of pins rotating around a vertical axis. Since the stimulus-figure provided no decisive information as to which set of pins was in front and which set was in back, the apparent depth-relation between the two sets of pins changed back and forth, and concomitant changes occurred in the apparent direction of rotation. These changes occurred so quickly and clearly that

some *Ss* thought, even after many measurements, that the changes were occurring in the figure itself. This figure has all the advantages of the Lissajous figure, but is superior to it in that rate of movement is stabilized by a constant-speed motor. Moreover, this figure is easier to calibrate in luminance, to maintain at a constant luminance, and to analyze in terms of its stimulus-properties.

Considerable interest attaches to the question of what constitutes the monocular cue to depth in the present ambiguous figure. It is planned to treat this question in detail in a later study, but it may be noted that the cue seems to be the change in lateral visual-angle velocity of a pin as a function of lateral position of the pin. Ames found that this type of cue did not give a striking depth-effect when a single stimulus-figure was used, but a strong depth-effect was obtained in the present study with the pattern of six pins.[7] It may also be noted that similar pin-patterns have been used previously by Metzger to study depth-effects resulting from movement.[8]

S's eyes were at the same level as the fixation-point and at an observation-distance of 24-in. An adjustable chin- and forehead-rest were provided to steady *S*'s head. An aperture in a horizontal slide was placed in front of his eyes, and stops were provided to locate this aperture before either the left or the right eye. It was desirable that a perpendicular from the fixation-point should coincide approximately with the visual axis of either the left or the right eye. To obtain these conditions, the entire box containing the stimulus was pivoted as shown in Fig. 1. Stops were provided for the extreme right and left positions, and these stops were adjusted for an interpupillary distance of 63 mm.

General method of measurement. The usual method of studying *RAC* is to measure the number of apparent changes during some specified period of observation, such as 1 min. Köhler, however, has published data from two *Ss* showing *RAC* during successive periods of observation.[9] He used a static ambiguous figure, and the periods of observation were separated by 1 min. Köhler found that the *RAC* increased with the number of observation-periods, although he found it impossible to make reliable continuous measurements because of the difficulties already mentioned with static figures. Philip and Fisichelli, working with Lissajous figures, also found an increase in *RAC* from the first to the second half of their 1-min. observation-period. These data suggested that if reliable measurements of *RAC* could be made as a continuous function of observation-time, the *RAC* would increase up to some maximal value. The apparatus developed for the present experiment made such measurements possible. This method was considered theoretically fruitful because it would yield the initial *RAC*, the slope of the increase, the amount of the increase, the maximal *RAC* attainable with prolonged observation, and the observation-time required for the maximal *RAC* to be attained. By comparison, the usual procedure would merely provide the sum over some arbitrary initial portion of this curve. The present experiment thus introduces the curve relating *RAC* to observation-time as a basic measure in studies of *RAC*.

Procedure. Eight *Ss* (3 men and 5 women) were used; all were between 18 and 29 yr. of age. There were no histories of injury or disease involving their eyes,

---

[7] W. H. Ittelson, *The Ames Demonstrations in Perception*, 1952, 64-66.

[8] Wolfgang Metzger, Tiefenerscheinungen in optischen Bewegungsfeldern, *Psychol. Forsch.*, 20, 1934, 195-260.

[9] Köhler, *Dynamics in Psychology*, 1940, 67-82.

visual pathways, or brains. Three *Ss* required glasses for the observation-distance of the experiment, and they wore their glasses while being tested. Six of the *Ss* did not know how the ambiguous figure was produced, or any of the purposes of the experiment. The other two *Ss* knew how the figure was produced, the purposes of the experiment, and the theoretically expected results. The results for these *Ss* showed no essential difference from the other results, however, so there was no indication that prior knowledge had any effect.

The same procedure was followed for all *Ss*. On the first day of measurement, *S* was given his instructions. He was told to fixate the small dot which would appear on the screen and to maintain fixation as steadily as possible during the two periods when the screen was illuminated. It was requested that a passive attitude be taken toward the figure at all times, and that no attempt be made to see any particular kind of movement. *S* was asked to keep track of the direction of movement of the pins which appeared to be in front, and to press a microswitch whenever the direction of rotation changed or whenever the kind of movement changed. Results were recorded on an Esterline-Angus recorder, and at the end of each test the results were tabulated by counting the number of apparent changes which occurred during successive ½-min. periods of observation. The *Ss* were not shown their results until the experiment was completed.

For each measurement, *S*'s eyes were first placed in correct position and the chin- and forehead-rests were adjusted. When making measurements with the right eye, the opening in the eye-selector was placed in front of the right eye and the stimulus-screen turned toward the right eye. When the left eye was tested, these positions were, of course, reversed. The pin-pattern was moved out of the projection-path, and the constant-speed motor was turned on. Then the room lights were turned off, the shutter was closed, and the point-source lamp was turned on. *E* gave a ready signal, and 2 sec. later opened the shutter to initiate a 1-min. preadaptation-period, during which *S* merely looked at the fixation-point. *E* gave another ready signal 2 sec. before the termination of the preadaptation-period. Then the shutter was closed for 1 sec. while the pin-pattern was drawn into place. When the shutter opened again, the recorder started simultaneously, and a 4-min. testing period was initiated. Preliminary measurements had indicated that the maximal *RAC* was achieved in that time.

The first day of measurement was only for practice, and two measurements were made with each eye. The eyes were tested in an alternating order, with a minimal separation of 30 min. between tests to prevent the increase in *RAC* during one test from carrying over to the next. The practice day was followed immediately by Day 1 and Day 2 of testing. On Day 1 the eyes were tested in a left-right-left-right order, with five measures in the morning and five in the afternoon. On Day 2 a right-left-right-left order was used. Thus 10 measurements were obtained on each eye, with order-effects counter-balanced. It should be noted particularly that the method of counter-balancing took care of any systematic change in *RAC* which may have occurred from morning to afternoon, or from Day 1 to Day 2.

RESULTS

The results of this experiment are shown in Table I, and the results for three of the *Ss* are also presented graphically in Fig. 2. These three indi-

vidual sets of results were chosen because they illustrate most of the findings of the experiment, and because they fit on a single graph without overlap.

All eight *S*s showed an increase in *RAC* as a function of observation-time. The general pattern of the increase may be seen in the curve of Fig. 3, which was fitted by inspection to average values for both left and right eyes of all *S*s combined (each value based on 160 measurements). This curve is a monotonic increasing function, which shows a striking negative acceleration. The *RAC* increases rapidly up to about 2 min. and

TABLE I

APPARENT CHANGES PER HALF-MINUTE AS A FUNCTION OF OBSERVATION-TIME FOR THE LEFT AND RIGHT EYES OF ALL EIGHT *S*s COMBINED

(Each value in the body of the table is the average of 10 measurements)

S	Eye	Observation-time in minutes								Av.
		0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	
AP	Right	3.5	3.3	3.5	3.7	4.2	4.9	4.2	4.1	3.92
	Left	3.5	3.5	2.9	4.2	4.3	4.7	4.3	4.0	3.92
AR	Right	3.7	3.7	4.4	3.7	4.3	5.3	4.5	4.8	4.30
	Left	3.9	4.3	3.2	4.3	5.0	4.7	4.0	4.8	4.28
CC	Right	3.8	3.1	3.8	3.7	4.6	5.4	5.0	4.9	4.29
	Left	3.9	4.0	5.0	5.1	6.2	4.9	5.0	4.7	4.85
EF	Right	1.3	0.6	1.5	1.5	2.1	2.5	2.7	2.5	1.84
	Left	1.0	0.6	1.6	1.6	1.8	2.1	2.2	2.6	1.69
HW	Right	3.1	5.8	6.1	8.4	7.6	7.0	6.1	7.7	6.48
	Left	3.9	5.3	5.6	5.3	6.9	6.7	5.7	6.9	5.79
JC	Right	2.4	3.3	3.7	3.8	4.5	4.7	4.7	5.1	4.02
	Left	2.5	3.4	3.7	4.4	4.8	5.0	5.1	4.5	4.18
LD	Right	2.4	2.4	3.5	3.8	4.4	5.1	5.1	5.2	3.99
	Left	2.1	3.1	4.0	4.4	4.4	4.6	4.7	3.9	3.90
VS	Right	1.4	1.5	1.8	2.3	1.9	2.1	2.6	2.8	2.05
	Left	1.5	1.6	1.7	1.7	2.1	2.5	2.0	2.2	1.91
Av.		2.74	3.09	3.50	3.87	4.32	4.51	4.24	4.42	3.84

then gradually levels off. Subsequent experiments have confirmed the general shape of this curve.

There was a positive but imperfect correlation in the present study between the initial *RAC* and the slope of the increase which occurred with prolonged observation. The nature of this correlation may be seen in Fig. 2. The initial *RAC*, however, also proved partially independent from the slope of the curve. Among the *S*s with a given initial *RAC*, there was considerable variation in the slope of the curve. Likewise, among the *S*s with a given slope, there was considerable variation in initial *RAC*. These findings are interpreted to mean that the physiological process which gives rise to apparent changes in an ambiguous figure is complex, involving a

minimum of two sub-processes. At least one sub-process is considered to affect both the intercept and slope of the curve, thus accounting for the general positive correlation noted; but at least one other sub-process is considered to affect only the intercept, or only the slope, thus accounting for the observed degree of independence of these two constants of the curve. In any case, the partial independence of the two constants means

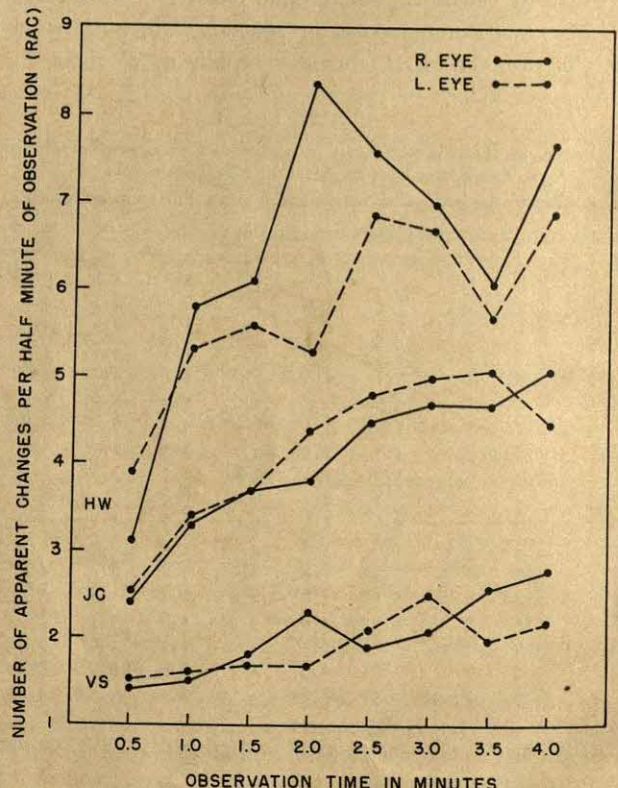


FIG. 2. *RAC* AS A FUNCTION OF OBSERVATION-TIME FOR THE LEFT AND RIGHT EYES OF THREE *Ss*

that curves showing *RAC* as a function of observation-time give more information than the usual measure of *RAC*.

The curves for the left and right eyes of a given *S* were very similar in five cases. In the data for three *Ss*, however, there was some indication that the curves for the left and right eyes might be slightly different. To determine whether any of these differences was significant, an analysis of variance was computed for the results of each *S*. The inhomogeneity of the within-cell variances, which increased roughly

with observation-time for all *Ss*, necessitated transformation of the raw data. A square-root transformation was used for all scores except those of *EF* and *VS*, for which the square root of X + 1 was taken. Cochran's test[9] of homogeneity was then applied to the transformed data, and no inhomogeneity of within-cell variances was indicated (0.05-level) in any of the eight cases. A separate analysis of variance was therefore performed on the transformed data of each *S*. The effect of observation-time proved significant at the 0.001-level in five *Ss*, at the 0.01-level in two *Ss*, and at the 0.10-level in one *S*. The difference between left and right eyes was significant at the 0.05-level only in the case of *CC*. On a purely chance basis, 1 out of 20 *Ss* would be expected to have a difference significant at the 0.05-level. Since 1 out of 8 was found in the present study, there was no convincing evidence for a true difference between the overall *RAC* of the left and right eyes in any of the *Ss*. The interaction between eyes and observation-time did not prove significant

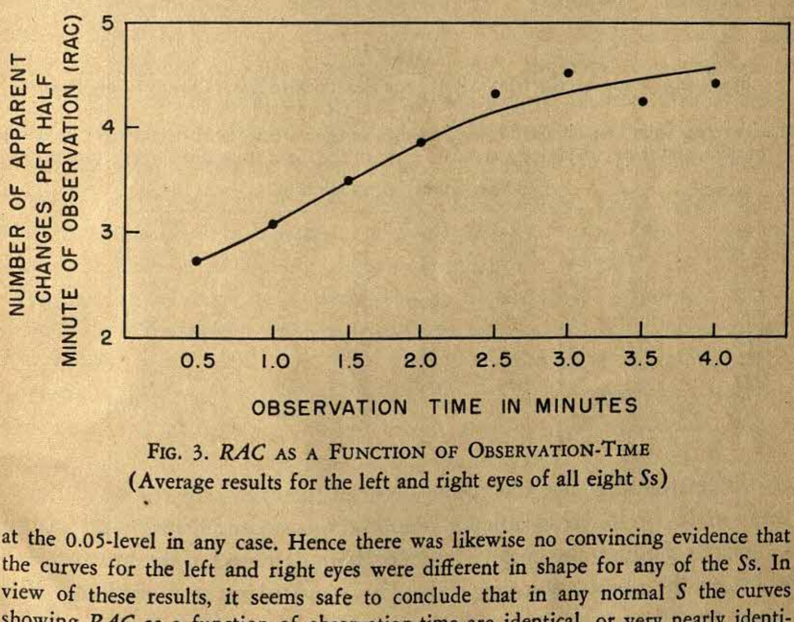


FIG. 3. *RAC* AS A FUNCTION OF OBSERVATION-TIME
(Average results for the left and right eyes of all eight *Ss*)

at the 0.05-level in any case. Hence there was likewise no convincing evidence that the curves for the left and right eyes were different in shape for any of the *Ss*. In view of these results, it seems safe to conclude that in any normal *S* the curves showing *RAC* as a function of observation-time are identical, or very nearly identical, in the left and right eyes. Mull, Arp, and Carlin determined the *RAC* of an outline cube over a 2-min. observation-period, and they also found no definite difference between the *RAC* of the left and right eyes of a given *S*.[10]

Individual differences proved quite large, as may be seen in Fig. 2. There were individual differences in the initial *RAC*, in the slope of the increase as a function of observation-time, in the magnitude of the in-

[10] H. K. Mull, K. Arp, and P. Carlin, Indications of a central factor in uncontrolled and controlled shifts in cube perspective, this JOURNAL, 65, 1952, 89-90.

crease as a function of observation-time, and in the maximal *RAC* attained with prolonged observation. There were also individual differences in the amount of observation-time required for the *RAC* to reach an approximate maximum.

The reliability of measurement in most *S*s proved quite good, as may be seen in the results of *VS* and *JC* in Fig. 2. The curves for these *S*s are relatively smooth, and the curves for the left and right eyes of each *S* agree well. The results of *HW* have also been included, however, to show the poorest reliability obtained. Since the present study was conducted partly to improve methodology for measuring *RAC*, computations were made of the error of measurement. The results for the left and right eyes of each *S* were pooled, and pooling was permissible because in a given *S*

TABLE II

Standard Error of the Mean *RAC* for Each *S* and Each Observation-Time (Left and Right Eyes Combined)

(Each average value was obtained by finding the average variance of the eight separate distributions, and then computing a standard-error of the mean from this average variance)

S	Observation-time in minutes								Av.
	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	
AP	.24	.31	.30	.27	.28	.29	.32	.29	.29
AR	.24	.29	.28	.26	.34	.37	.35	.44	.36
CC	.23	.26	.39	.40	.43	.31	.36	.31	.34
EF	.15	.13	.17	.14	.25	.26	.21	.28	.21
HW	.34	.44	.63	.73	.73	.65	.45	.51	.58
JC	.21	.24	.25	.29	.34	.27	.34	.40	.30
LD	.20	.25	.32	.29	.40	.36	.36	.43	.34
VS	.14	.11	.12	.19	.22	.24	.35	.21	.21
Av.	.23	.27	.34	.36	.40	.36	.35	.37	

no systematic difference was found between the eyes with respect either to mean *RAC* or to error-variance. Pooling was accomplished with each *S* by merely finding the mean *RAC* for both left and right eyes at each point on the observation-time axis. The standard error of each of these means was then computed, and these standard errors are shown in Table II. The average standard error of the mean has also been computed for both columns and rows in Table II, to show changes with observation-time and among *S*s. It may be noted that the standard error of the mean varied directly with observation-time, roughly paralleling the increase in *RAC*. The standard error of the mean likewise varied among *S*s, and these differences were also directly related to differences in *RAC*. Since the regression of standard error of the mean on mean *RAC* proved to be linear, a product-moment correlation was computed between the standard errors in

the body of Table II and the means with which they are associated. The value obtained was 0.87, and the limits which include the true value at the 0.01-level are 0.76 to 0.93. Thus there is a strong relation between mean *RAC* and the standard error of the mean.

In addition to the finding that measurements in most *S*s are highly reliable, there are preliminary results which indicate that the curve for any given *S* is relatively stable over periods of at least several months. A number of *S*s have been tested repeatedly over periods of from three to six months, and the curve for each *S* has been found relatively stable over these periods of time.

Discussion

In general, the type of ambiguous figure used in this experiment proved satisfactory. The only important difficulty was the occurrence of a considerable amount of extraneous movement. Extraneous movement is defined here as any kind of apparent movement other than smooth rotary movement in one direction or the other, and it may be defined more generally as any kind of apparent movement other than the two primary types of movement in an ambiguous figure. Philip and Fisichelli also found a considerable amount of extraneous movement with Lissajous figures. Hence it is difficult, if not impossible, to avoid. The kind of extraneous movement most often seen with the present figure was a two-dimensional movement in which the end-pins would both move toward the middle, bump into each other, and move back to their end-positions. Sometimes the pins would also weave in and out between one another. Still another kind of apparent movement consisted of two sets of rotating pins standing side by side, one set of pins on either side of the fixation-point. When this kind of apparent movement occurred, the apparent direction of rotation in the two sets of pins proved to be independent. In a few cases the pins on the extreme right and left were reported to stand still, while the four center pins revolved. One *S* also reported the extreme right and left pins to stand still, while the left-center pair and right-center pair revolved independently. Thus a number of kinds of extraneous movement were reported with this ambiguous figure. Since the occurrence of extraneous movement increases the difficulty of interpreting results, and since it may also have an adverse effect upon reliability of measurement, efforts are now being made to find ambiguous figures in which the frequency of extraneous movement is reduced.

The frequency of occurrence of extraneous movement varied widely among *S*s. Some *S*s never saw extraneous movement, but *HW* estimated

that he saw such movement about 80% of the time. Furthermore, extraneous movement was seen more frequently as observation-time increased. Fisichelli noted that these same two factors are important in the frequency of extraneous movement in Lissajous figures.[11]

If the increased *RAC* as a function of observation-time were not entirely dissipated by the beginning of the next measurement, the initial *RAC* of the second measurement would be increased. The data were analyzed for such an effect, but it was not found. Thus the 30-min. interval between tests was sufficient to prevent the increase in *RAC* during one test from carrying over to the next. Since blinking sometimes seems to 'trigger' an apparent change, a control experiment was performed in which *RAC*-curves and blink-rate curves were recorded simultaneously. In the control experiment the *RAC* increased with observation-time, but there was no increase in blink-rate. Hence blinking had no appreciable influence on the shape of the *RAC*-curve.

All findings of the present experiment are consistent with the hypothesis that the same physiological process is responsible for both figural after-effects and apparent changes in an ambiguous figure. Two findings are particularly noteworthy in this regard. First, the general shape of the curve showing *RAC* as a function of observation-time is like that of the curve determined by Hammer for figural after-effects as a function of inspection-time.[12] Secondly, the *RAC* curves for left and right eyes of a given *S* proved essentially identical. Köhler and Wallach report that the physiological process responsible for figural after-effects is located at a cortical level. Hence if apparent changes in an ambiguous figure are caused by the same physiological process, the *RAC* curves for left and right eyes of any given *S* should be virtually identical. Of course these two findings do not provide crucial evidence for the hypothesis that the same physiological process is responsible for both phenomena, but the findings are consistent with the hypothesis. Fisichelli noted that with Lissajous figures the *RAC* seems to be directly related to the amount of stimulation per unit time and area which is delivered to the retina by the Lissajous figure.[13] Since *RAC* and the figural after-effect are both a function of stimulation, Fisichelli considers that the findings with Lissajous figures also support the hypothesis.

---

[11] Fisichelli, Reversible perspective in Lissajous figures; some theoretical considerations, this JOURNAL, 60, 1947, 240-249.

[12] E. R. Hammer, Temporal factors in figural after-effects, this JOURNAL, 62, 1949, 337-354.

[13] Fisichelli, *op. cit.*, 240-249.

If the same physiological process is responsible for both phenomena, it seems certain that the *RAC*-curve provides an improved method for measuring that process. Most published measurements of figural after-effects have contained serious methodological errors.[14] Even when these errors have been eliminated, however, the reliability of measurement has been quite low.[15] The author has made a careful attempt to develop a reliable method for measuring figural after-effects, but little success has been attained. The basic difficulty is that the effects are quite small relative to the error of measurement which must be tolerated. Thus satisfactory measurements of figural after-effects are at least difficult to obtain, and measurements of figural after-effects as a function of inspection-time are particularly difficult and laborious. The present experiment indicates clearly, however, that curves showing *RAC* as a function of observation-time can be obtained with relative ease and a high degree of reliability.

SUMMARY

An improved method of measuring rate of apparent change (*RAC*) in a dynamic ambiguous figure is described. When *RAC* is plotted against observation-time, a monotonic increasing function with a definite negative acceleration is found. The *RAC* increases rapidly up to 2 min. and then gradually levels off. Marked individual differences appear in initial *RAC*, in the rate of increase with observation-time, in the amount of increase with observation-time, in the maximal *RAC*, and in the amount of observation-time required for the curve to reach an approximate maximum. The initial *RAC* is partially independent of the increase which occurs with prolonged observation. Such a curve therefore yields more information than does the usual measure of *RAC*. The curves for the left and right eyes are identical, or very nearly identical, for each of the eight *S*s measured. All findings of the present experiment are consistent with the hypothesis that the same physiological process is responsible for both figural after-effects and apparent changes in an ambiguous figure. If this hypothesis is correct, curves showing *RAC* as a function of observation-time provide a better method than the figural after-effect for measuring that process.

[14] K. T. Brown, Methodology for studying figural after-effects and practice effects in the Müller-Lyer illusion, this JOURNAL, 66, 1953, 629-634.

[15] Hammer, *op. cit.*, 337-354.

PARALLAX AND PERSPECTIVE DURING AIRCRAFT LANDINGS

By JAMES J. GIBSON, PAUL OLUM, and FRANK ROSENBLATT,
Cornell University

The performance of successfully landing an airplane is generally supposed to be dependent on the ability to perceive visual depth and distance, or more generally *space*. The classical explanation of human ability to see space is that the eye provides a set of *cues* to the distance of a given object, and likewise differential cues to the different distances of different objects. Such cues include (a) binocular parallax or disparity, (b) linear perspective, including the apparent size of objects whose real size is known, (c) aerial perspective, superposition, shading, and (d) motion parallax, *i.e.* the apparent movement of physically stationary objects relative to one another and to the observer. The last depends upon motion of the observer, usually locomotion, and the strength of the effect is in proportion to the velocity of the observer. Consequently it ought to be of particular significance for spatial judgments during aircraft landings, in which locomotion occurs at a considerabe velocity. The aim of this report is: (1) to examine the existing definitions of motion parallax; (2) to show that they are not sufficiently general to cover the visual situation during aircraft landing; (3) to formulate geometrically a somewhat different principle (called motion perspective); and (4) to apply this general principle to the problem of perceiving the ground and one's relation to it during descent.

DEFINITIONS OF MOTION PARALLAX

The term 'parallax' comes from astronomy. It means the difference in the direction of a body, that is its apparent displacement, caused by a difference in the position of the observer (*O*). The parallax of a star, for instance, can be obtained by measuring its apparent displacement relative to more distant stars when the earth is to the right of the sun (in midwinter) and to the left of the sun (in midsummer). From its parallax, the distance of the star can be computed. A parallax of one second of arc corresponds to a distance of one 'parsec,' which is 200,000 times the distance to the sun.[1]

The term 'motion parallax' or 'monocular motion parallax' came into use when men began to study visual space perception, and its meaning is derived from the

* Accepted for publication July 31, 1954. This research was supported in part by the United States Air Force under Contract No. AF33(038)-22373, monitored by the Skill Components Research Laboratory, Air Force Personnel and Training Research Center.

[1] F. W. Dyson, Parallax, *Encycl. Brittan.*, 17, 1929, 261-262.

astronomical one. It refers to the apparent motion of a near object relative to a far object, this being to the left whenever *O* moves to the right and to the right whenever *O* moves to the left. This relative apparent motion in the visual field probably became obvious when railroad trains made rapid locomotion a commonplace. In Boring's historical survey, he could find no discussion of this visual criterion of distance prior to the time of Helmholtz; perhaps, as he suggests, because it was considered self evident.[2] Helmholtz himself described the sensation of near objects gliding past us as we advance through the countryside, of more distanct objects doing so more slowly, and of remote bodies maintaining their permanent position in the field, and then he added: "Evidently under these circumstances the apparent angular velocities of objects in the field of view will be inversely proportional to their real distances away; and consequently safe conclusions can be drawn as to the real distance of the body."[3] This statement seems to define what is meant by motion parallax as clearly as any to be found in the psychological literature. The phenomenon has been given rather scant consideration; Carr's book on space perception fails even to mention it.[4] Tschermak, however, has argued that it ranks with binocular parallax in importance, the two principles being similar.[5]

Very recently, Graham has defined the cue of motion parallax by the principle that whenever an *O* moves in the environment, "a differential angular velocity exists between the line of sight to a fixated object and the line of sight to any other object in the visual field."[6] The differential motions of the corresponding images lead to statements by *O* such as that near objects move 'against' the direction of his own movement and that far objects move with his own movement. Graham has demonstrated that visual acuity for such differential motions is very great.

Insufficiency of Motion Parallax During Aircraft Landings

It should be noted that the apparent relative motion of bodies during *O*'s movement depends on his being displaced to right or left of the line between him and the bodies in question. Such a lateral displacement does always occur for the planetary movement of the earth in space, but it does not occur in the same way for a linear movement in terrestrial locomotion. Motion parallax is very evident when the passenger in a railway train looks out of the side window but it is not so obvious when the locomotive engineer looks forward along the tracks. Above all, lateral displacement is not sufficient to describe the movement of the flier during a landing. In short, *motion parallax does not occur in the way so far defined for objects ahead during forward locomotion.*

If the movement of the earth in astronomical space were linear, and if all other

---

[2] E. G. Boring, *Sensation and Perception in the History of Experimental Psychology,* 1942, 263-308.

[3] Hermann von Helmholtz, *Physiological Optics,* vol. 3, 1925, 295.

[4] H. A. Carr, *An Introduction to Space Perception,* 1935, 1-413.

[5] Armin Tschermak-Seysenegg, Ueber Parallaktoskopie, *Arch. f. d. ges. Physiol.,* 241, 1939, 454-469.

[6] C. H. Graham, Visual perception, in S. S. Stevens (ed.), *Handbook of Experimental Psychology,* 1951. It may be noted here, to anticipate a more general principle, that this assertion should be modified by adding the qualification *except when both objects are in the line of movement of the observer.* The importance of this modification will become evident.

bodies were stationary, their apparent displacements would be a function not only of their distance but also of their angular deviation from the line of movement of the earth. Crudely this could be expressed by saying that the apparent movement of the heavenly bodies would be a sort of radial expansion from an apex in the sky at the point toward which the earth was moving.[7]

Evidently Helmholtz's rule for drawing 'safe conclusions' about the real distance of a body in space is not so safe after all since it is not sufficient. Apparent angular velocity is inversely proportional to distance, true enough, but it is also a function of direction from *O*'s line of locomotion. As angular deviation from this line decreases, apparent motion also decreases, and finally vanishes. In other words, the apparent velocities of objects scattered about the visual field seem to be complicated by the fact that they 'carry information' not only about the distances of the objects but also about the direction in which *O* himself is moving.

Phenomenon of Motion Perspective

We are concerned with the aircraft landing situation, however, and this is not one in which there are objects scattered about an otherwise empty visual field; it is one in which there is a continuous extended surface, the earth. The senior writer has argued that, in general, the practical problems of space-perception are actually problems of surface-perception.[8] The geometry we should employ for their analysis must take account of the presence of the surface. If differential velocities for different parts of the retinal image constitute a basis for distance-perception during locomotion, as we have every right to expect, they should be analyzed as a gradient of velocities in the visual field corresponding to the elements of a physical surface. A beginning of this task has been made by the senior writer in the description of what is called *motion perspective*.[9] Diagrams have been presented to illustrate the differential velocities at different points of the earth's surface for various fields of view relative to the line of *O*'s movement. Each diagram, representing velocities by vectors, is a projection of the ground on a 'picture plane' in front of the eyes.[10] They have received a sort of empirical verification by the analysis of motion picture shots of the ground taken from an airplane in different directions relative to the line of flight.[11]

For example, during straight and level flight with the eyes fixed on the horizon, the ground ahead seems to expand radially from a center on the horizon; the ground to right or left seems to be skewed; and the ground

[7] J. J. Gibson, *The Perception of the Visual World*, 1951, 123.
[8] Idem.
[9] Idem., ch. 7.
[10] Idem, Figs. 53-58 in ch. 7.
[11] Gibson (ed.), *Motion Picture Testing and Research*, AAF Aviation Psychology Research Report No. 7, 1947, 230 ff.

behind seems to contract radially toward another center on the horizon. All velocities vanish at the horizon. During a landing glide, the projection of the ground seems to expand radially from a center at the intersection of the glide-line with the ground, to reach a maximum between that point and the horizon, and again to vanish at the horizon.

It may be noted from these diagrams that in the case of a plane surface parallel to the line of locomotion, the apparent velocities in the visual field of the points on any line of direction from *O* to the horizon are actually inversely related to the distances of the corresponding points on the surface. This fact is superficially similar to the rule of Helmholtz concerning motion parallax. For locomotion *not* parallel to a plane surface, the fact, however, no longer holds. The apparent velocity of spots on a surface not only diminishes toward zero at the horizon of the surface; it also diminishes toward zero at the spot where the line of locomotion intersects the surface. The velocity-distribution over the projection of a 'longitudinal' surface is negatively correlated with the distribution over the surface of distances-from-here, but the velocity-distribution for a 'frontal' surface is not. In the latter case, the gradients of velocity starting with the horizon and ending with the nearest point of the surface are complicated, as it were, by the center of radial expansion which is no longer on the horizon. The utility of this focus for various problems of landing an airplane has been described elsewhere,[12] and will be further discussed in a future report.

The question arises, then, whether the retinal image of a frontal surface contains 'information' about the differential distances of the spots of the surface, and if so what variable of the deforming image carries this information. Does the pattern or distribution of angular velocities in the rays of light reflected from a surface to a moving point determine that surface? What information does the pattern of velocities include and is it sufficient by itself (without regard to binocular vision or ordinary perspective) to enable the flier to make a landing?

Mathematical Analysis of Motion Perspective

The description of motion perspective so far given is not geometrical but phenomenological; it is an introspective account of the visual field of an *O* during locomotion. The visual field is the way the world looks when it is seen not as a world but as a picture, and the illustrations previ-

[12] Idem, 228 f.

ously published have been projections on the plane of a picture. Such a description of optical sensations is useful but it does not take the place of a mathematical description of optical stimulation, *i.e.* the retinal image, or preferably the sheaf of light-rays entering the eye. An exact statement of motion perspective is required. For this purpose we need to consider

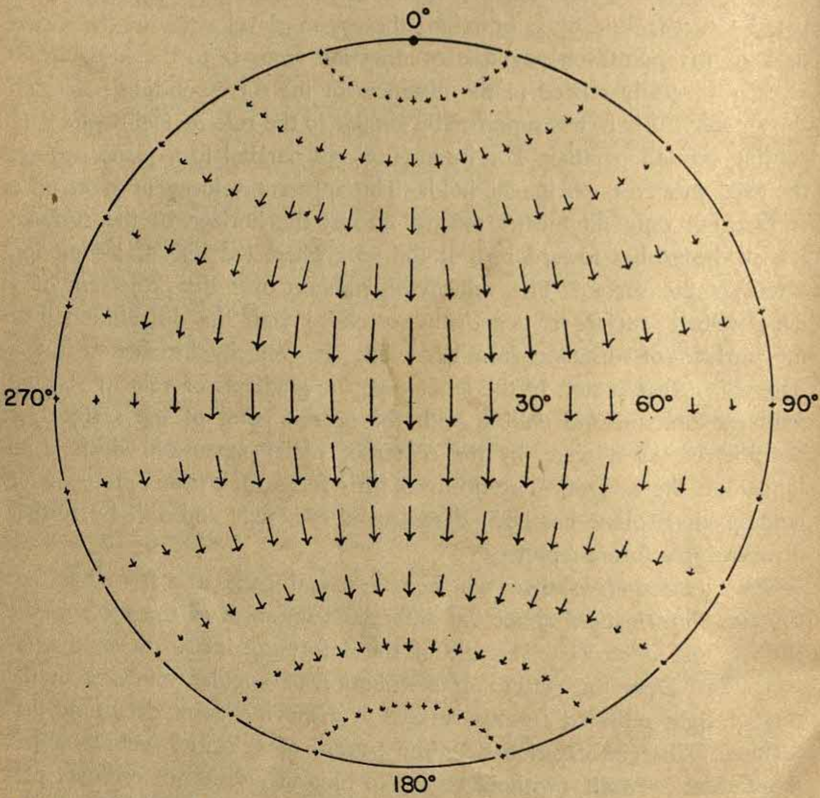


FIG. 1. PROJECTED DIFFERENTIAL VELOCITIES IN LEVEL FLIGHT FROM A LONGITUDINAL SURFACE PARALLEL TO THE LINE OF LOCOMOTION
In this figure distance from the center represents the angle δ' of Fig. 3 and the angle around the circle represents the angle θ' of Fig. 3.

a purely abstract mathematical situation in which an eye is reduced to a geometrical point, the earth to a plane, the texture of the earth to points on the plane, and the light reflected from the earth to lines intersecting in the point.

The situation to be considered, then, consists of a point moving in a given direction with respect to a plane surface. We will assume (1) that the

plane has 'elements,' (2) that the elements reflect light, each element being indicated by a ray, and (3) that the surface is so large that it reflects to the eye a sheaf of rays approaching 180° in angular extent. Optically speaking, the surface has a horizon in any direction from the eye. The point (eye) may move parallel to the plane, or at some angle of

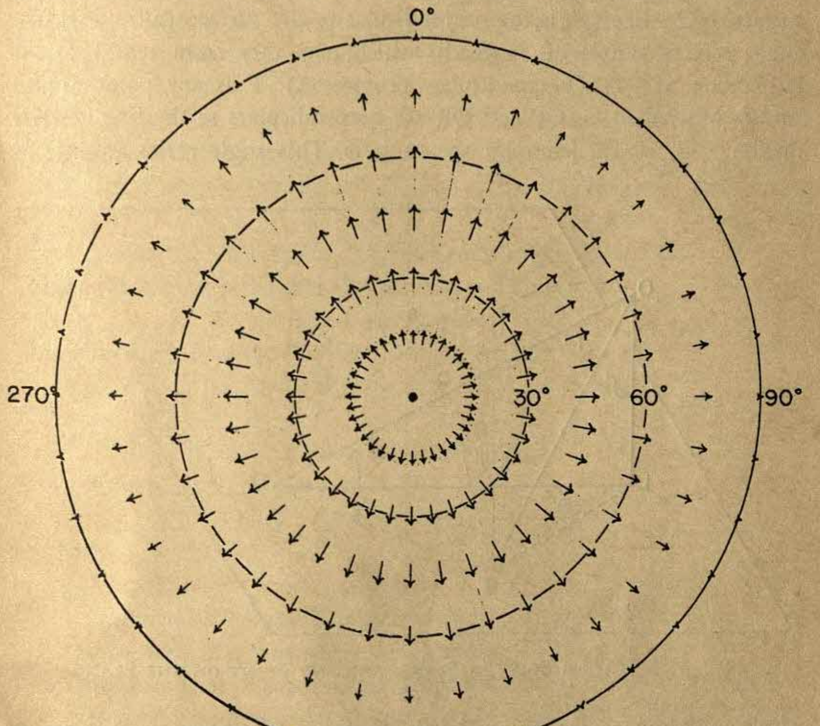


FIG. 2. PROJECTED DIFFERENTIAL VELOCITIES IN A HELICOPTER LANDING FROM A FRONTAL SURFACE PERPENDICULAR TO THE LINE OF LOCOMOTION

Distance from the center represents the angle δ' of Fig. 3 and the angle around the circle represents the angle θ' of Fig. 3.

inclination to it, or perpendicular to it. When the plane is taken to be the earth, these three possibilities represent the visual situation during, respectively, ordinary locomotion or level flight, an airplane landing, and a vertical helicopter landing.

We need make no assumptions as to the direction in which the eye is fixated or whether it undergoes a rotary eye-movement. We do not assume

anything about gravity or the posture of *O* relative to gravity. The surface considered could be any surface (*e.g.* a wall or ceiling as well as a floor). We are concerned merely with the distribution of angular velocities in the sheaf of rays reflected to a moving point from a large surface.

Fig. 3 shows the abstract situation. *O* is the observer (the eye or, more exactly, its nodal point) moving toward the point G on the ground (the surface). The line OL is the perpendicular to the surface (altitude). The angle β is the angle of approach, which may vary from zero (parallel locomotion to 90° (perpendicular locomotion). P is any point on the surface other than G. QP and QR are perpendiculars to the line of locomotion, OG, which determine the angle Θ. This angle varies around the

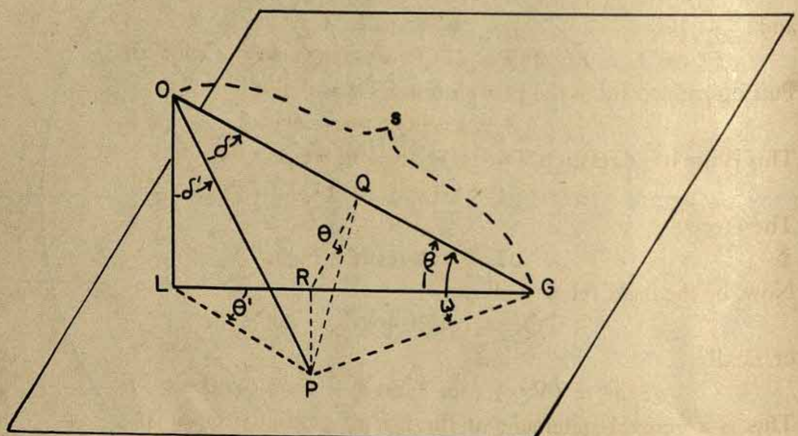


FIG. 3. A MOVING POINT, A PLANE, AND ANY POINT ON THE PLANE

clock according as the point P falls to the right, above, to the left, or below the line of locomotion. The angle δ is the angular separation of the light rays reflected from P and from G. As *O* approaches the ground this angle increases, that is, the ray from G is motionless but the ray from P moves away from it (hence the radial apparent motion of P, and the apparent 'expansion' of the ground outward from G).

Note that the angle Θ remains fixed as *O* approaches the ground. Consequently the apparent expansion of the ground at the point P will have the magnitude $d\delta/dt$ and will be along the direction of constant Θ. We shall call the 'expansion vector' at *P* then a vector tangent to a line $\Theta =$ constant and of magnitude $d\delta/dt$.

The general description of motion perspective then requires a mathe-

matical computation of this expansion vector for all points P and all angles β of approach. The linear velocity of approach V, or $-ds/dt$, is assumed to be constant; s refers to the distance OG in Fig. 3. As coördinates to describe the point P we shall use the pair of angles (Θ, δ) or later, for purposes of graphical representation, the pair (Θ', δ') of Fig. 3.

In Fig. 3, it should be noted that, since QP and QR are perpendicular to OG, PR is also perpendicular to OG. PR is perpendicular then to LG and to RQ. In short the diagram should be viewed as a set of adjoining right triangles. The angle ω is merely an auxiliary variable which is useful in the derivation of the general formula. We have

$$\cot \delta = OQ/PQ = (s - PQ\cot\omega)/PQ = (s/PQ) - \cot\omega,$$

and

$$\cos \Theta = RQ/PQ = (PQ \cot\omega \tan\beta)/PQ = \cot\omega \tan\beta.$$

Putting the second in the first eliminates ω and gives:

$$\cot \delta = (s/PQ) - \cos \Theta \cot\beta.$$

This is the basic relation. Differentiating it, we get

$$-\csc^2\delta\,(d\delta/dt) = (1/PQ)\,(ds/dt).$$

Therefore

$$d\delta/dt = V\,(\sin^2\delta/PQ).$$

Now, by the basic relation above

$$1/PQ = (\cot \delta + \cos \Theta \cot \beta)/s$$

or finally

$$d\delta/dt = (V/s)\,(\sin \delta \cos \delta + \sin^2\delta \cos \Theta \cot \beta). \quad \ldots\ldots [1]$$

This is a general statement of the law of motion perspective.[13]

It can be seen that when β is 90° (the case of the helicopter landing) the expression after the plus sign drops out and the pattern of angular velocities is simply a function of $\sin \delta \cos \delta$, which yields the symmetrical radial pattern of Fig. 2 when plotted on angular coördinates.

To get the pattern for the case $\beta = O$ (the case of level flight) we replace s by $h \csc \beta$ where $h = OL$, the altitude, and let β go to zero.[14] This gives then

$$d\delta/dt = (V/h)\,(\sin^2 \delta \cos \Theta). \quad \ldots\ldots\ldots\ldots [2]$$

and yields the pattern of Fig. 4 when plotted on angular coördinates.

---

[13] The above law can be generalized to describe the apparent motion of stationary bodies in empty space by adding one constant to the equation. If the point P is taken to be a point on *any* plane, not merely points on the plane of Fig. 3, and if Z is taken as the distance between its plane and the reference-plane of Fig. 3, then the formula becomes $d\delta/dt = (V/(s - Z \csc\beta)\,(\sin \delta \cos \delta + \sin^2\delta \cos \Theta \cot \beta)$.

[14] The general equation then becomes $d\delta/de = (V/h)\,(\sin \beta \sin \delta \cos \delta + \cos \beta \sin^2 \delta \cos \Theta)$.

This plot represents the angular velocities of points on the ground only when δ is less than 90°. That is to say, it represents only the ground 'ahead' during level flight, and the empty half of the plot represents the sky.

When β is 45° (which would be the case of a very steep glide) the pattern of angular velocities is one in which the velocities vanish at the horizon of the surface combined with another in which the velocities vanish at the focus of radial expansion. To plot this pattern on an easily comprehensible circular graph, it is preferable to make the center of the angular coördinate system not at the focus of radial expansion but at the

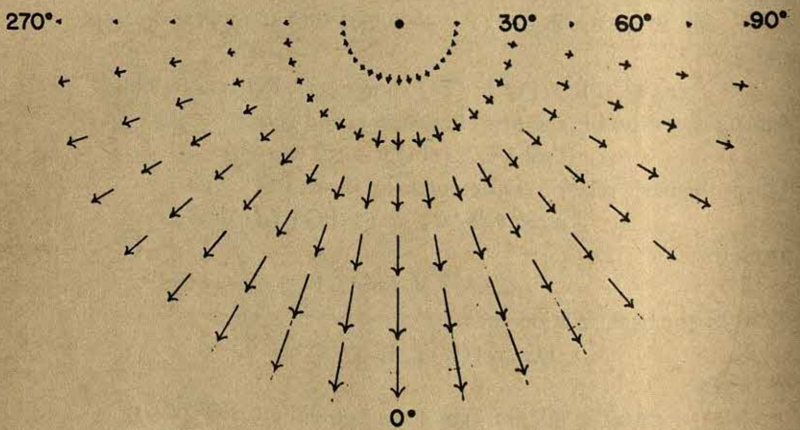


FIG. 4. THE PATTERN OF VELOCITIES DURING LEVEL FLIGHT
This is the same as Fig. 1 except that here distance from the center represents the angle δ of Fig 3 (only the range from 0° to 90° is shown) and the angle around the circle represents the angle Θ of Fig. 3.

center of the sheaf of rays reflected from the surface, that is, at the nearest point of the surface. For this purpose the formula must be converted. We need to express the angular velocity of δ in terms not of δ and Θ but of δ' and Θ' in Fig. 3. It may be noted that $\delta' = \delta$, and $\Theta' = 180° - \Theta$ when β is 90°. The converted formula is

$$d\delta/dt = (V/h) \cos^2 \delta' \left[(\cos \beta - \sin \beta \tan \delta' \cos \Theta')^2 + (\tan \delta' \sin \Theta')^2\right]^{1/2} \quad [3]$$

Using this formula, the pattern of velocities for a 45° approach is plotted in Fig. 5. It represents a view looking down at the earth, like Fig. 2, as distinguished from a view looking toward the horizon, like Fig. 4.

The formula just stated can also be used to plot the pattern of angular

velocities looking down at the earth when β is zero (the case of level flight). This plot is given in Fig. 1. Figs. 1, 2, and 5 are all angular coördinate plots of a sheaf of light rays to a point from an infinitely large plane surface, when the point is moving in different directions relative to the surface.

We conclude that the motion perspective of a surface like the earth, or a

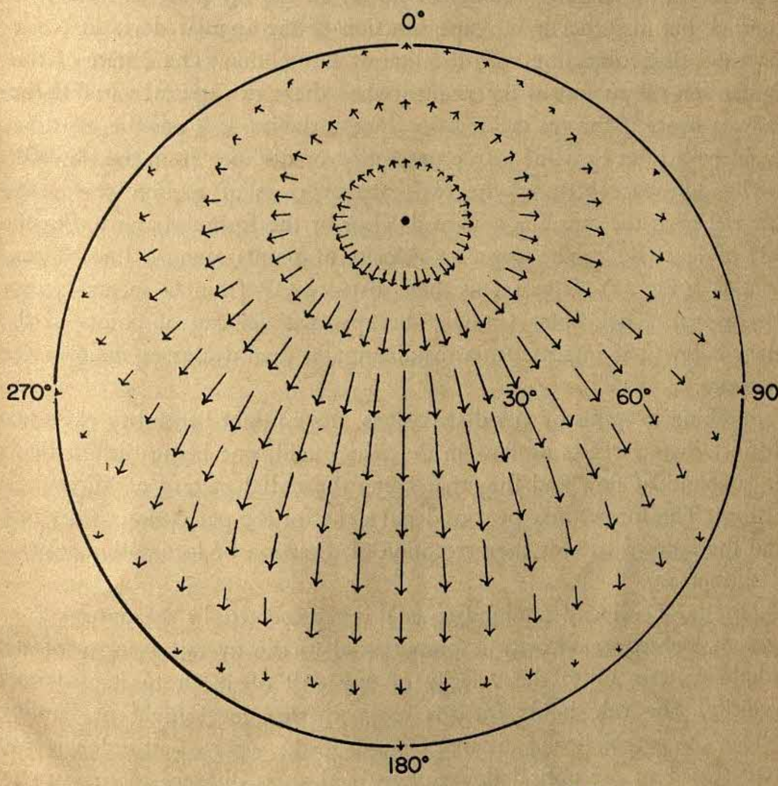


FIG. 5. THE PATTERN OF VELOCITIES IN A 45° GLIDE
Distance from the center represents the angle δ′ of Fig. 3 and the angle around the circle represents the angle Θ′ of Fig. 3. Note that the arrows point along lines of constant Θ.

floor or wall, carries information about the direction of one's locomotion (the angle of approach to the surface) as well as a great deal of information about the surface itself.

The line OP in Fig. 3 is of considerable significance. It is the distance from *O* to any point on the ground, and it is for this distance that motion

parallax (the relative angular velocity) is classically supposed to provide the cue. It follows at once from Equation [1] that

$$ds/dt = \mathrm{V}(\sin\,\delta/\mathrm{OP})$$

and hence that

$$\mathrm{OP} = \mathrm{V}(\sin\delta)/(d\delta/dt).$$

This says that the distance of any point on the surface is inversely proportional to the angular velocity of its ray in the ray-sheaf as Helmholtz asserted, but also that it is a sine function of the angular deviation of its ray from that coinciding with the line of locomotion. The distance of an element on the surface is determined when these two optical variables are known; hence if the eye can register these variables, it is possible that they might be a joint cause of a true perception of distance from the observer.

The difference between the two extreme cases of motion perspective referred to at the outset is now made clear by the formulas above. For the case of level flight, the apparent velocity of points along a line of constant Θ' (Fig. 3) decreases as their distance OP from *O* increases. For the case of a helicopter landing, the apparent velocity of points in the plane below first increases to a maximum and then decreases again as *OP* increases.

Speaking in terms of visual sensations, there might be said to exist two distinct characteristics of flow in the visual field, one being the gradients of 'amount' of flow and the other being the radial pattern of 'directions' of flow. The former may be considered a cue for the perception of distance and the latter a cue for the perception of direction of locomotion relative to the surface.

O's linear velocity (ground speed) is represented in the optical flow-pattern. Subjective velocity is proportional to the overall velocity of the whole pattern, or to the velocity of any part of it, or to its maximal velocity. The rule holds for any case, whether level flight or landing glide. The perpendicular distance from *O* to the surface (altitude) is also represented in the optical flow-pattern, and so is distance to the surface on the line of locomotion in the case of a landing glide. Both are inversely related to its velocity. Ground-speed and altitude are not, however, independently determined by the optical information. A more rapid flow-pattern may indicate either an increase in speed or a decrease in altitude. Length of time before touching down, however, seems to be given by the optical information in a univocal manner.

The above formulation of motion perspective for a plane surface implies that a plane surface is uniquely specified by the formula. A curved surface,

convex or concave, is excluded. In other words the optical stimulation for a moving *O* indicates the shape of the surface in depth from which the light is reflected. It does not merely provide a cue to the depth of the surface if by that is meant information of limited statistical reliability; it contains the depth of the surface as a variable of the array. Evidently, the traditional assumption of visual theory that rays of light cannot specify the distance from which they come is misleading. They can at least specify an important property, the curvature, of the surface from which they are reflected. In this sense *it is possible to assume that the retina can register the displacements in space of the eye itself.*

All these considerations suggest that the psychology of aircraft landing does not consist of the classical problems of space perception and the cues to depth. It is a psychology of locomotion, which occurs in time as well as in space, and the problems are those of the *judgments required for the control of locomotion.* The fundamental visual perception is that of *approach to a surface.* This percept always has a subjective component as well as an objective component, *i.e.* it specfies *O*'s position, movement, and direction as much as it specifies the location, slant, and shape of the surface. The variables, or discriminable features, of this percept are probably ones like altitude and speed, direction of glide, distance and time on the present path to the ground, and distance and time on any other path to the ground. Along with these go variables of the flatness of the ground (or slant or curvature), the size or scale of its parts, and the direction of its parts. The critical discriminations required for landing are probably just these.

It must be remembered that the mathematical formula given above did not take any account of the 'texture perspective' of the surface considered. The latter was necessarily assumed to have elements but no assumption was made that these were of equal size, or were equally spaced, or even that equal areas of the surface tended to have equal numbers of elements. If any of these additional assumptions are made (and they are applicable to actual surface in our environment, such as landing fields) an additional type of perspective exists—the perspective of size and density.[15] Human optical stimulation, moreover, usually includes a third type of perspective —that of binocular disparity of retinal images. If all three of these are considered together, for the situation of a binocular *O* moving in relation to a textured surface, the information about the properties of the

[15] Gibson, *The Perception of the Visual World,* 1951, ch. 6.

surface contained in the optical stimulation is multiply assured and plentiful. So also is the information about the position and motion of *O* relative to the surface. The performance of landing an airplane depends on these types of information. The principle of motion perspective is the same whether the information depends upon daylight, moonlight, landing lights, flares, beacons, or radio waves.

It should also be remembered that this formula for motion perspective does not take into account stimulation produced by the pull of gravity on *O*'s body (or by any other force such as acceleration). The direction of locomotion *relative to the surface* is given, parallel, perpendicular, or inclined, but the direction of locomotion relative to the line of gravity (or a resultant vector) is not given by visual stimulation.[16]

There is probably a psychological connection between the perception of locomotion relative to a surface and the perception of the surface itself as stationary and rigid. Motion perspective as a kind of experience has been described by terms like 'apparent expansion' of the surface. A more important experience, however, is the correct one of a stationary and perfectly rigid surface. It would be easy to assume that the first kind of experience causes the second—that sensations cause perceptions. Another possible assumption would be that the two kinds of experience are alternatives and that when the optical deformation is not a stimulus for the impression of an expanding scene it necessarily becomes a stimulus for the impression of locomotion in a rigid world. A still better assumption might be that the correct perception is determined by certain invariant properties of the stimulus-array, and that what causes the impression of an 'expanding' scene is a complicated matter of no great importance. A discussion of these hypotheses will not be continued here, although they are important for psychological theory.

Another omission is the analysis of *O*'s retinal image when eye-movements of pursuit or change in fixation complicate the stimulus-situation. This must be left for a future report. There is reason to believe, however, that the fundamental relations of motion perspective are preserved in the retinal image during eye-rotations; the facts here presented are unaffected by them.

SUMMARY

The accepted formulation of monocular motion parallax as a cue to distance is shown to be insufficient, particularly when applied to the prob-

---

[16] Gibson, The relation between visual and postural determinants of the phenomenal vertical, *Psychol. Rev.*, 59, 1952, 370-375.

lem of locomotion relative to a surface such as the earth. A more general description of the parallax phenomenon is given under the name 'motion perspective.' The latter holds for a *O*'s movement along any line parallel or inclined to any extended surface, which includes the case of aircraft landings.

A mathematical analysis of motion perspective is presented in terms of the optical flow-pattern reflected from a surface to an eye. It is shown that the variables of this flow-pattern are specific not only to the 'depth' of the surface but also to the movement of an *O*. Assuming that these variables are stimuli for perception, they can determine not only the experience of a stable tridimensional world, but provide a basis for the judgments required for the control of locomotion in that world.

COMPARATIVE OLFACTORY INTENSITIES OF THE ALIPHATIC ALCOHOLS IN MAN

By Lawrence Kruger, Alvin N. Feldzamen, and Walter R. Miles, Yale University

Olfaction remains the most neglected field of sensory physiology today. Men and insects have been the usual subjects, and the quantitative data are largely confined to liminal measurements. The special technical problems involved in olfactometry and the failure of threshold data to provide a reliable index of supra-threshold intensity have recently led to the development of a method of matching olfactory intensities analogous to that long used in visual and auditory studies.[1] A scale for the quantitative determination of supra-threshold olfactory intensity has been proposed.[2]

In the present paper we report the results of measurements of olfactory intensity for a homologous series of aliphatic alcohols. A homologous series was chosen and studied because chemical differences between the odorants would be minimized and olfactory differences more readily related to changes in physical properties. This emphasis is in accord with current opinion that physical rather than chemical processes are critical in olfactory stimulation.[3]

It is hoped that comparison of the olfactory effectiveness of various odorants with their physico-chemical properties will reveal information relevant to the physiological mechanisms of olfaction.

Materials and Methods

The method used in this study is based on intensive matchings. A dilution series of n-heptanal in benzyl benzoate constituted the standard scale against which other

* Accepted for publication September 14, 1954. This study was supported by grants-in-aid from Arthur D. Little, Inc., and the Research Corporation, New York City. The authors wish to express their thanks to Dr. Lloyd H. Beck, Department of Psychology, for his advice and encouragement and to Drs. Lyons and Sturtevant of the Sterling Chemistry Laboratory, Yale University, for providing space and guidance for the measurement of vapor pressures.

[1] B. M. Wenzel, Techniques in olfactometry: A critical review of the last one hundred years, *Psychol. Bull.*, 45, 1948, 231-247.

[2] L. H. Beck, Lawrence Kruger and P. Calabresi, Observations on olfactory intensity. I. Training procedures, methods and data for two aliphatic homologous series. *Ann. N. Y. Acad. Sci.*, 58, art. 2, 1954, 225-238; Lawrence Kruger, A. N. Feldzamen, and W. R. Miles, A scale for measuring supra-threshold olfactory intensity, this Journal, 68, 1955, 117-123.

[3] Beck and Miles, Some theoretical and experimental relationships between infrared absorption and olfaction, *Science*, 106, 1947, 511; Miles and Beck, Infrared absorption in field studies of olfaction in honeybees, *Proc. Nat. Acad. Sci.*, 35, 1949, 292-310.

odorants were compared. One milliliter of each odorant, standard or unknown, was placed in a 10 × 75 mm. test tube, from which the *S*s sniffed. Upon being presented with an unknown, *S*, by making sniff-comparisons, endeavored to find the tube in the standard scale most closely approximating the overall intensity of the unknown. He was encouraged when he thought it appropriate to interpolate the unknown between two standard tubes.

The unknowns were the 10 normal aliphatic alcohols, as follows: C_3, 1-propanol; C_4, 1-butanol; C_5, 1-pentanol; C_6, 1-hexanol; C_7, 1-heptanol; C_8, 1-octanol; C_9, 1-nonanol; C_{10}, 1-decanol; C_{11}, 1-undecanol; C_{12}, 1-duodecanol. Determinations were made 5 min. or more apart to obviate the effects of olfactory adaptation. The entire series of 10 alcohols was presented in varying orders on 8 to 10 separate days to every *S*. The unknowns in the first series were all saturated aliphatic alcohols. In the second series some of these alcohols were diluted with benzyl benzoate, as will be described below. Three extensively trained *S*s made the comparisons which constitute the quantitative data presented in this paper.

The standard scale is a geometric dilution series, constructed so that tube No. 1 is 100% n-heptanal and tube *n* is $100\%/2^{n-1}$ n-heptanal in benzyl benzoate by volume, with the strongest intensity No. 1 and successively weaker odors proceeding to No. 9. Intensive values are represented by tube numbers. A description of the scaling-procedure and more complete physical and chemical units for the scale, together with details of the training procedure and the sniffing method, are presented elsewhere.[4] It should be recognized that in this, as in all human olfactory studies, absolute effective stimulus-concentrations cannot at present be stated because the actual stimuli are conducted through convoluted passages and through interfaces of undetermined character. Perhaps as a consequence of this, the scale-values obtained for one substance by two different *S*s are often dissimilar, and, to compare closely different substances attention must generally be restricted to the matching-judgments of one individual.

The chemicals used in this study were obtained from the Yale Sterling Chemistry Laboratory and from Polak's Frutal Works, Amersfoort, Holland, and their purity was insured by redistillation. The infra-red spectra were obtained on a Perkin-Elmer infra-red spectrophotometer, model 21, using liquid samples in sodium chloride cells. The solvent used was carbon tetrachloride.

RESULTS

As individual differences are to be expected in olfactory studies we present no group averages. In Fig. 1 is represented the intensity scale positions of the saturated aliphatic alcohols from 1-propanol to 1-dodecanol for 3 *S*s separately. Every point plotted represents the mean of at least eight matching judgments and the length of the vertical line above or below the center of a point represents an estimated standard deviation. *S*s 1 and 3 were Caucasians with light brown hair, *S* 2 was an Oriental. There is an apparent trend toward decreased intensity from about C_5 to C_{12}. *S* 2 shows a fairly

[4] Kruger, Feldzamen, and Miles, *op. cit.*, 117-123.

regular decline of intensity from C_3 to C_{12}, with the exception of C_{11}. *S* 3 demonstrates similar results, with the exception that C_3 was judged to be weakest in intensity of the 10 unknowns. For *S* 1 there was a steady increase from C_3 to and including C_6. Beyond C_6, which was judged the most intense of the 10 samples, the trend was toward a decreasing intensity, with the exception of C_{11}. All the aliphatic alcohols in saturated solution were

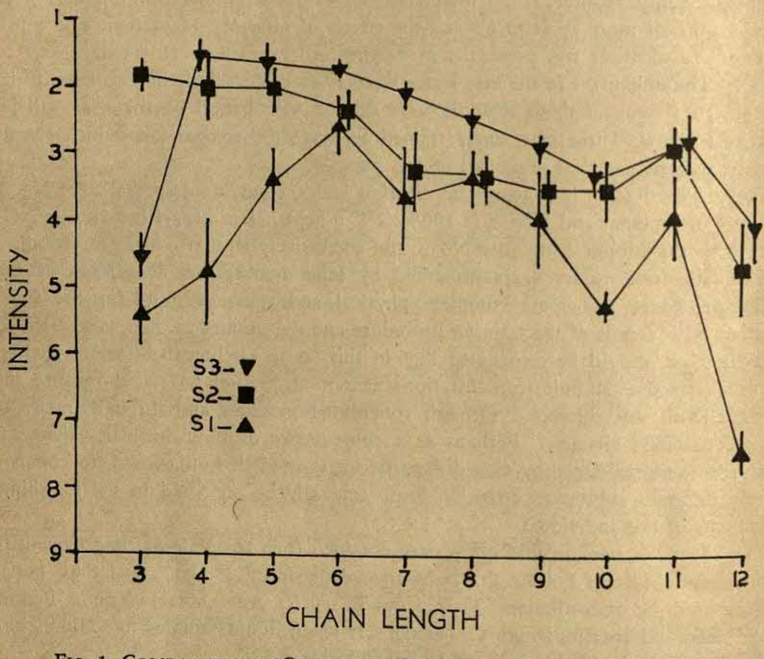


FIG. 1. COMPARISON OF OLFACTORY INTENSITY WITH CHAIN LENGTH FOR 10 NORMAL ALIPHATIC ALCOHOLS

judged by all three *S*s to give less intense odors than 100% n-heptanal (scale intensity No. 1). The more intense alcohols were judged equivalent to scale Nos. 1.5–2.5. Each of the alcohols except C_3 was given a lower scale number by *S* 2 than by *S* 3; however, three of these nine differences were 0.5 or less. *S* 1 gave all the aliphatic alcohols lower intensity ratings than *S* 3, and with the exception of C_8 his ratings were also lower than found by *S* 2. The consistent rise of intensity found by all 3 *S*s at C_{11} seems a genuine olfactory phenomenon, probably due to instability of this substance as a consequence of its high rate of oxidation. Methyl and ethyl alcohols were not studied in this investigation because these compounds,

as well as other substances of high vapor pressure, are nasal irritants, which *S*s do not sniff uniformly. It is well-known, however, that these alcohols have higher thresholds[5] and, in saturated solutions, are less intense as odors than the alcohols studied in this experiment.[6]

To evaluate the effect of decreased concentration, diluted series (with benzyl benzoate) of unknowns of the alcohols from 1-butanol to 1-dodeca-

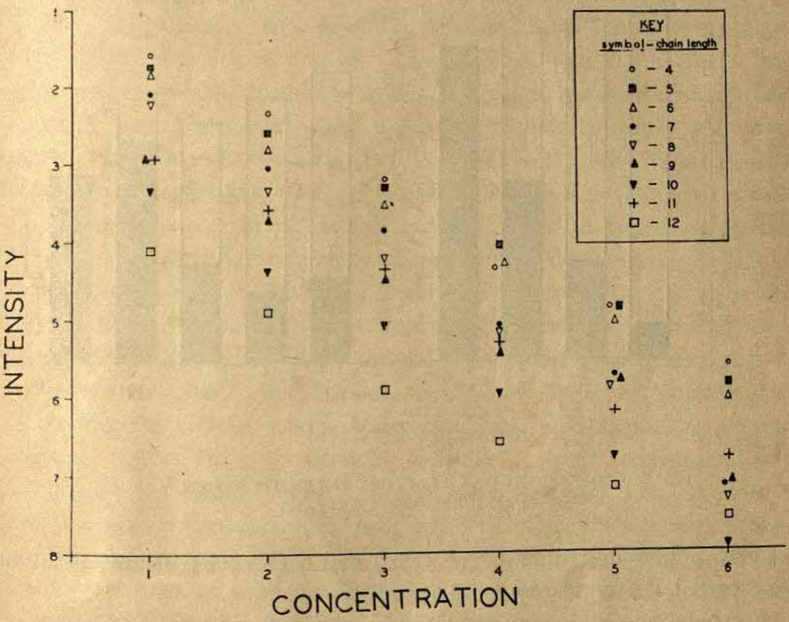


FIG. 2. COMPARISON OF OLFACTORY INTENSITY WITH CONCENTRATION FOR NINE NORMAL ALIPHATIC ALCOHOLS BY *S* 3

nol were constructed, and the solutions were matched against the standard scale. The variations with concentration in these nine alcohols for *S* 3 are presented in Fig. 2. The concentration units were the same for both the standard scale and the alcohol solutions. The alcohol concentration range used was 100% to 3.12% thus making a series of six unknowns for each of the nine alcohols.

Taking into account the probable error of judgment, we see throughout this wide range of concentration that the relation between intensity and

[5] J. Passy, L'odeur dans la série des alcools, *Compt. Rend Soc. Biol.*, 44, 1892, 447-449.

[6] E. L. Backman, Note sur la puissance des odeurs et leur solubilité dans l'eau et dans l'huile, *J. Physiol. et Path. Gen.*, 17, 1917, 1-4.

chain length for these alcohols approaches stability. Furthermore, intensity is approximately a linear function of concentration. The maintenance of this consistency at low intensities suggests that threshold measurements should show the same trend toward decreasing effectiveness of stimuli with increasing chain length. In Fig. 2 we again find C_{11} judged more intense than

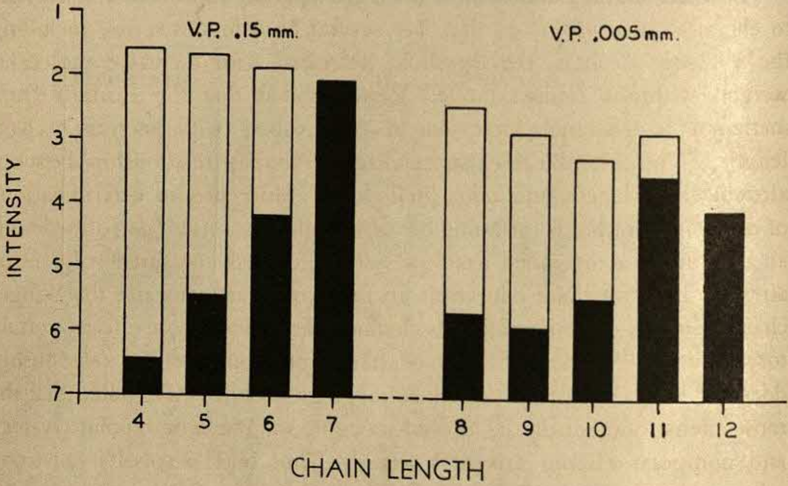


FIG. 3. COMPARISON OF OLFACTORY INTENSITY WHEN VAPOR PRESSURES ARE CONSTANT

C_9 or C_{10}. Low intensities—5 (6.25%) and 6 (3.12%)—show, as might be expected, less regularity.

Computation of olfactory intensity in terms of equal numbers of molecules in the nose is undoubtedly of more theoretical significance than consideration of the customary concentration units for solutions. An approximate demonstration of the truth of this statement can be had by comparing solutions at the same vapor pressure. Those concentrations of the alcohols from 1-octanol to 1-undecanol having the same vapor pressures at 20° C as 1-dodecanol (approximately 0.005 mm. Hg) were determined, and, from Fig. 2, the intensities at these concentrations were extrapolated. For the shorter chain alcohols these concentrations were too low to permit intensity extrapolation, so a similar computation was made using the vapor pressure at 20° C of 1-heptanol (approximately 0.15 mm. Hg). In Fig. 3, the top of the dark portion of each vertical bar represents the rated intensity corrected for vapor pressure; the top of the entire bar represents the intensity of the undiluted alcohol without correction for vapor pressure. In

these calculations, ideal behavior of both solutions and gases was assumed throughout. In terms of equal vapor pressure, the longer chain lengths show stronger odor intensities.

Discussion

The early investigators who studied the relation of olfactory sensitivity to chemical structure noted that, for several homologous series, including the aliphatic alcohols, the threshold decreased with increased molecular weight, within a limited range.[7] Passy claimed that the alcohols from methanol to pentanol increased in effectiveness with increased chain length;[8] Von Hornbostel also reported a similar relationship between alcohol chain length and odor 'helligkeit.'[9] More precise determinations of olfactory thresholds for a number of homologous series (performed with an odiometer, a controlled gas-flow system in which measured volumes of air were used to dilute odorants) also seemed to demonstrate that longer chain members of an homologous chemical series were more effective olfactory excitants.[10] Backman, however, has pointed out that this relationship does not exist throughout the entire range of alcohols.[11] He noted that the most intense odorants in the alcohol series are soluble in both polar (water) and non-polar (lipoid solvents) liquids. Thus butyl alcohol is a strong odorant; methyl and ethyl alcohols, insoluble in lipoid solvents, and cetyl alcohol (C_{16}), lipoid-soluble and water-insoluble, are low in odor intensity. Because chemical differences generally outweigh solubility properties, the rule of lipoid and water solubility must in comparisons be confined to a series of similar substances. In a study of various paraffin derivatives, including the aliphatic alcohols from methyl to amyl, Cook[12] found that their stimulative effectiveness for insects increased with decreased vapor pressure.[13]

In more recent studies of contact chemoreception and olfactory stimula-

---

[7] W. J. Crozier, Chemoreception, in C. Murchison (ed.), *Handbook of General Experimental Psychology,* 1934, 987-1076; E. G. Boring, *Sensation and Perception in the History of Experimental Psychology,* 1947, 447; R. W. Moncrieff, *The Chemical Senses,* 1951, 279.

[8] Passy, *op. cit.,* 447.

[9] E. M. von Hornbostel, Ueber Geruchshelligkeit, *Arch. f. d. ges. Physiol.,* 227, 1931, 517-538.

[10] V. C. Allison and S. H. Katz, An investigation of stenches and odors for industrial purposes, *J. Ind. Eng. Chem.,* 11, 1919, 336-338.

[11] Backman, *op. cit.*

[12] W. C. Cook, The effectiveness of certain paraffin derivatives in attracting flies, *J. Agri. Research,* 32, 1926, 347-358.

[13] *Op. cit.,* 356.

tion in blowflies, Dethier and his co-workers observed that the members of homologous aliphatic series (including the alcohols) are rejected at logarithmically decreasing concentrations as the carbon chain length is increased.[14] Although the taste-stimulating efficiency increases directly with chain length in both insects and man,[15] our Figs. 1 and 2, and the left-hand portion of Fig. 3, indicate that the reverse relation occurs in human olfaction, as reported by our trained *S*s.

The discrepancy between the results of insect and man may be due to several factors. The olfactory organs of insects differ markedly in structure from the vertebrate end-organs. Threshold data and suprathreshold intensity may not be related in any simple fashion. Many substances detected in minimal concentrations are not intense odorants when their concentration is increased. It is unlikely, however, that the discrepancies between threshold data and supra-threshold intensity in the alcohol series can be explained in this manner, since the same intensity relationships persist throughout a wide range of dilutions (Fig. 2) and because thresholds of rejection and acceptance in insects show the same trend.[16] We cannot now compare intensive measurements with threshold determinations in man throughout the range of alcohols explored in this study, because sufficient threshold data are not available.

Another attempt to rate the supra-threshold intensity in the range of alcohols considered here was made by Crocker and Henderson.[17] These workers assigned arbitrary number values to odorants, denoting intensities ranging from 1 to 9, in each of four primary qualities: fragrant, acid, caprylic, and burnt. This subjective approach is difficult to relate to the present study. Summation of the four Crocker-Henderson numbers for each alcohol produces a series of numerical values which show essentially the same trend in intensity as found in the present study. Even the curious rise in intensity found for 1-undecanol in the present study is reproduced. This rise at C_{11} has also been noted in an acetate ester series.

The general decrease in intensity with increasing chain length (Fig. 1)

---

[14] V. G. Dethier and L. E. Chadwick, Rejection thresholds of the blowfly for a series of aliphatic alcohols, *J. Gen. Physiol.*, 30, 1947, 247-253; V. G. Dethier and M. T. Yost, Olfactory stimulation of blowflies by homologous alcohols, *ibid.*, 35, 1952, 823-839; Dethier, The physiology of olfaction in insects, *Ann. N. Y. Acad. Sci.*, 58, art. 2, 1954, 139-157; Dethier, in K. D. Roeder (ed.), *Chemoreception in Insect Physiology*, 1953, 544-576.

[15] Dethier, Taste sensitivity to compounds of homologous series, *Amer. J. Physiol.*, 165, 1951, 247-250.

[16] Dethier, *op. cit.*, *Ann. N.Y. Acad. Sci.*, 58, 1954, 149.

[17] E. C. Crocker and L. F. Henderson, Analysis and classification of odors, *Amer. Perf. Essent. Dil. Rev.*, 22, 1927, 3-11.

might have been predicted from the vapor-pressure decline with increasing chain length. The equal vapor-pressure graph (Fig. 3) indicates that this relation is not a molecular phenomenon. The assumption of ideality in the computation leading to this figure may not be warranted at these low concentrations, and consequently these data probably deserve less confidence

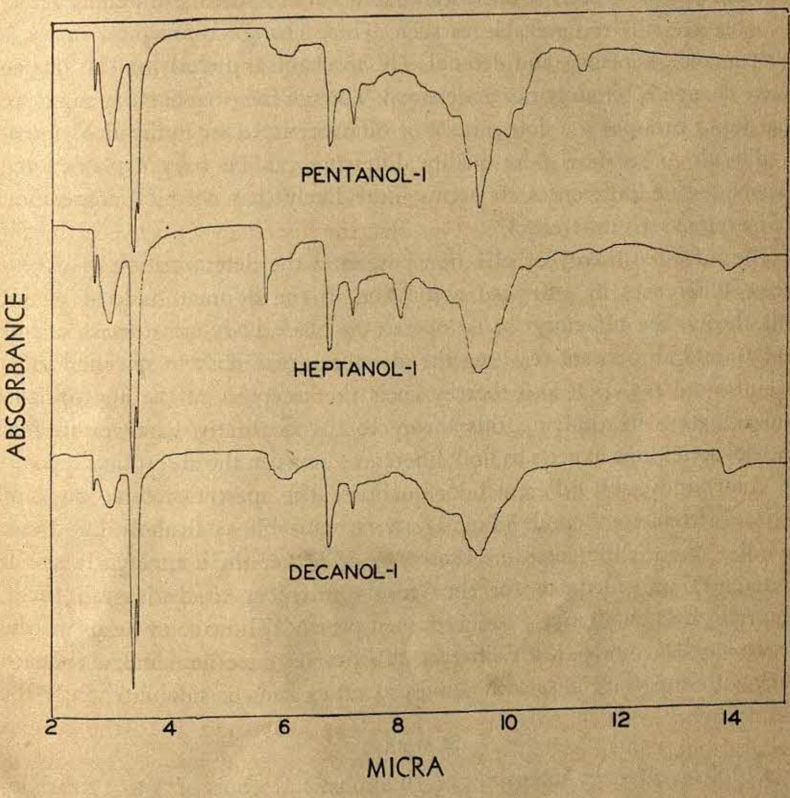


FIG. 4. INFRA-RED ABSORPTION SPECTRA OF SAMPLES OF THE THREE ALCOHOLS WHOSE ODORS WERE DISTINCTIVE

than the strictly formal concentration-intensity relations of Fig 2. It would seem, however, that the vapor-pressure decline with increasing molecular weight is more than sufficient to explain the decrease in intensity of these alcohols used in concentrated form. Were a constant vapor-pressure graph flat, it could be concluded that vapor pressure alone accounted entirely for the intensity change; this might tend to support a 'chemical' theory of olfaction. The apparent molecule-for-molecule increase in intensity, coin-

ciding with the same relationship occurring for taste, indicates that longer molecules are more efficient stimulants.[18] In view of the prevailing diversity of opinion with regard to the mechanism of olfactory stimulation this is no more easily explained than the reverse relationship.

In the course of this study, each of the *S*s reported that certain changes in odor quality appear as one ascends the alcohol series, although all the alcohols are still recognizable as such. These changes were most marked for pentanol, heptanol, and decanol. The mechanism underlying this qualitative change is certainly not understood. Though the present study suggests that vapor pressure is a determinant of olfactory intensity in the alcohols, it is difficult to see how odor-quality difference can be fully explained by vapor-pressure differences. It seems more likely that chemical impurities may explain such differences.

The infra-red theory of olfaction postulates the determination of qualitative differences by infra-red absorption in the odorant molecule.[19] In this theory, the olfactory 'hairs' operate as black body-radiators at about 37° C. and an odorant reaching the olfactory organ absorbs specific bands of infra-red radiation and thereby cools the receptors at the appropriate wave-lengths. In applying this theory to the qualitative changes in the alcohol series, one expects to find differences between the absorption spectra of compounds with different odor qualities. The spectra of three qualitatively different compounds (Fig. 4) are recognizable as alcohols, but there are also minor differences in absorbency. Whether these apparently small differences are adequate for the sensory receptors to distinguish these three alcohols qualitatively is uncertain at present. Those compounds whose chain lengths coincide with different qualities, on the other hand, also coincide with important physicochemical properties, such as solubility.

Summary

(1) The olfactory intensities of 10 aliphatic alcohols of 3 to 12 carbon chain lengths have been determined against a standard intensity-scale in a matching procedure by 3 *S*s trained in olfactory observation.

(2) The alcohols which are soluble in both polar and non-polar solvents were found to have the more intense odors.

(3) Within most of the range of alcohols studied it was found that olfactory intensity decreases with longer chain molecules. This trend appears, however, to be reversed after correcting for vapor-pressure.

[18] Dethieo and Chadwick, *op. cit.,* 247.
[19] Beck and Miles, *op. cit.,* 511.

(4) The same relative intensities in comparison with the standard scale were found at six different dilutions of each alcohol.

(5) The infra-red absorption spectra of 3 alcohols which were reported to be qualitatively distinctive were found to be similar, although slight differences were noted.

(6) Individual variations are quite distinct despite reasonably good consistency in the results for each *S*.

GENERALIZATION IN SPATIAL LEARNING

By W. LYNN BROWN and C. E. HUMPHREY, University of Texas

When discussing 'place learning,' Woodworth asks what does an animal learn, and answers his own question by stating that "the most obvious answer . . . is simply . . . the place." By 'place' he means "a concrete situation containing objects in spatial relation."[1] Woodworth's contention was experimentally confirmed in 1944 by Blodgett and McCutchan who showed that "there is an area some 24 in. in diameter within which all points may be thought of as having spatial equivalence. In this area, so far as response is concerned, the rat would be helpless if it were called upon to master a serial problem in which two or more choice-points were . . . 12 in. apart."[2]

The Blodgett-McCutchan study was the beginning of a long series of researches, of which the present study is a part. The first of the subsequent series was reported by Tolman, Ritchie, and Kalish in 1946. They found in confirmation of Blodgett and McCutchan's results that "rats do learn to expect goals in specific locations."[3] In a study conducted at the same time, Brown found to the contrary that they do not. Using a 12-cul tunnel-maze which was so arranged that the 'start' and 'finish' were in the same box, Brown found that his rats, when fed on a test-trial in the starting-box before they ran the maze, refused to eat until the maze had been run.[4] The conflict between the results of these two studies led to the study by Gentry, Brown, and Kaplan in 1947.[5] They too were unable to confirm the Blodgett-McCutchan results. Since their failure may have been due to their method, the problem was again attacked by Gentry, Brown, and Lee in 1948.[6] Negative results were again obtained. Although conditions were favorable during the preliminary training for learning the location of the goal and for developing a directional orientation with reference to it, they found no evidence that their rats had acquired even the rudiments of a general directional orientation. They concluded therefore that their findings and those of Gentry, Brown, and Kaplan "constituted a body of

---

* Accepted for publication July 15, 1954. This investigation was supported in part by a grant from the Research Institute of The University of Texas.

[1] R. S. Woodworth, *Experimental Psychology*, 1938, 135.

[2] H. C. Blodgett and Kenneth McCutchan, Choice-point behavior in the white rat as influenced by spatial opposition and by preceding maze sequence, *J. Comp. Psychol.*, 37, 1944, 66.

[3] E. C. Tolman, B. F. Ritchie, and Donald Kalish, Studies in spatial learning: I. Orientation and the short-cut, *J. Exper. Psychol.*, 36, 1946, 24.

[4] W. L. Brown, The effects of a common start-finish locus on orientation and behavior in a multiple T-maze, *J. Comp. Psychol.*, 39, 1946, 331-338.

[5] George Gentry, W. L. Brown, and S. J. Kaplan, An experimental analysis of the spatial location hypothesis in learning, *J. Comp. & Physiol. Psychol.*, 40, 1947, 309-322.

[6] George Gentry, W. L. Brown, and Hun Lee, Spatial location in the learning of a multiple T-maze, *J. Comp. & Physiol. Psychol.*, 41, 1948, 312-318.

evidence casting serious doubt on the validity of the assumptions underlying the spatial location-hypothesis."[7] Despite all of this evidence to the contrary, the conclusion drawn by Blodgett and McCutchan, and by Tolman, Ritchie, and Kalish, has in general been accepted.

The present study was designed specifically to answer the following questions which were involved in the preceding studies: (1) Does spatial learning in the white rat constitute the learning of a definite location in space? (2) Does it require a particular concrete situation? (3) Can it be generalized? (4) If it is not the learning of a finite location and if generalization does occur, upon what factor or factors does it depend? (5) Are specific, extra-maze, visual cues necessary?

METHOD

Apparatus. An elevated T-maze, 7 ft. 6 in. long and 2 in. wide, bisected by a 2-ft. starting-strip of similar width, was used in this study. It was elevated 15 in.

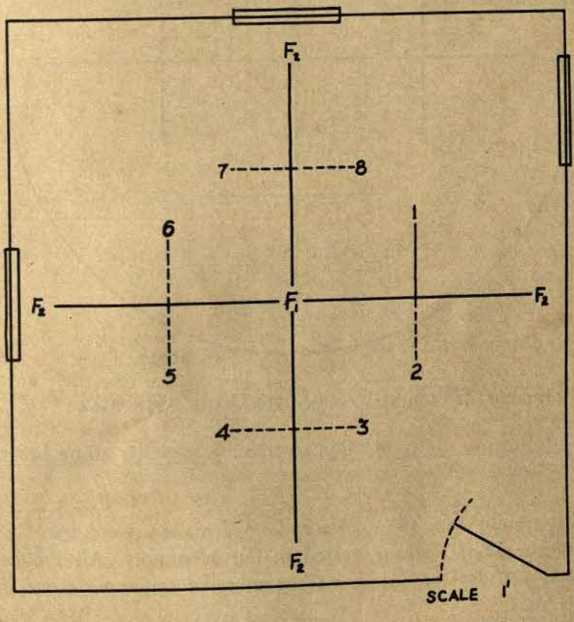


FIG. 1. FLOOR PLAN OF SQUARE ROOM WITH MAZE

from the floor. The pattern of the maze is shown in the eight positions in which it was used in Fig. 1. Only one section, however, was set up for a given trial.

The maze was painted with glossy black enamel. Two mazes of the same dimen-

[7] *Idem,* 318.

sions and patterns were used. One was placed in a square room with windows in the west, north, and east walls, and the other in an experimental dome (see Fig. 2). A food dish was fastened to the end and below the level of the maze that it would not be seen by the rat from afar. It was in the geometric center of the room (F_1) during all trials made by Group C and by Groups 1 and 2, and at the wall-end of the maze (F_2) for Group 3.

To avoid the possibility of tracking by smell, the large unit of the maze was rotated 180° after every run. The entire maze was washed daily with hot water

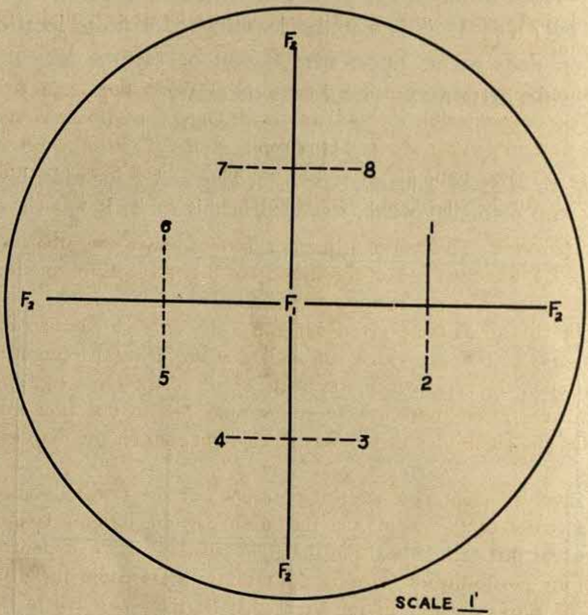


FIG. 2. FLOOR PLAN OF THE DOME WITH MAZE

and repainted following the completion of training for each group before test-trials were given.

Subjects. Sixty-eight rats (43 males and 25 females) were used in the study. They were of the Wistar strain, reared in our laboratory. All the animals were between 90 and 120 days of age at the beginning of the experiment, and all were experimentally naïve.

The *Ss* were divided into 4 groups of 17 each: experimental groups (Groups 1, 2, 3) and a control group (Group C). Four preliminary runs were given each group daily for five days on a straight elevated pathway. The preliminary trials were given in the colony room, not in the experimental rooms. Moist food was used in all phases of the experiment.

Experimental rooms. A square room (18 ft. × 18 ft., with a 10.5-ft. ceiling)

was used for training Group C and Groups 1 and 3, and for test-training Group 2. The floor was painted dull gray and the walls, ceiling, woodwork, and hardware were painted a flat ivory just before the study began. One window is centered in the north wall and one in the west wall. An off-center window and a locked door are in the east wall. Access to the room is by the door in the south wall, as shown in Fig. 1.

The experimental dome (18 ft. in diameter with 5.5-ft. side walls arched to an 8-ft. center) was used in the training of Group 2 and the test-training of Group C. It is constructed symmetrically of wooden ribs and covered with a fine mesh screen painted silver gray, hence gives a homogeneous surface radially symmetrical about the geometric center.

Scoring. An error was recorded if *S* went to or crossed a line 12 in. from the end of the cul. The noncorrection method was used. Only one error was counted if *S* made an incorrect run. After a run, *S* was replaced at the starting-position and this procedure was repeated until a correct run was made. Each *S* was permitted to eat at the food dish for 20 sec. upon completion of a correct run.

Procedure: Group C. To answer a question formulated in the problem, a control group (Group C) was used to test the hypothesis that place-learning is the learning of a definite location in space, hence generalization to other locations in space could not occur. The control *S*s were given one run a day in each position of the maze during seven days of training—each run ending at an identical location denoted by the food dish upon the maze which was at the center of the room (see Fig. 1). The training runs were given consecutively at Positions 1–8 to test later for response-disposition. On the eighth day each *S* was given eight runs on the diagonal positions shown in Fig. 3.

This test involved eight new starting-positions but the reward-location was the same—at the center of the room. On the ninth day the training conditions were maintained except that each animal ran the eight positions in a randomized sequence to determine any possibility of learning the problem by response-disposition, *i.e.* by the learning of turn-sequence. In the event that learning was due to this type of disposition, a randomized run should result in chance performance. The following day a completely new maze was used to determine the effects, if any, of intra-maze cues. The *S*s were run on Positions 1–8 as in training.

Group C was then transferred to the dome for seven days of training on the maze used by Group 2 (see Fig. 2). This procedure was designed as a test of generalization from one concrete situation to another concrete situation. The dome was lighted by a 7.5-w. white frosted bulb centrally located in the ceiling at a height of 7 ft. from the floor. A circular disk 12 in. in diameter was suspended below the bulb to reflect the light to the homogeneous ceiling.

Group 1. Group 1 was used to study the possibility of generalization from one axis in the room to three other axes in the same room. For the first seven days this group ran from Positions 1 and 2 only (see Fig. 1). *S*s were alternated between Positions 1 and 2 in a randomized order. On the eighth day, seven days of training began on Positions 1–8, with the first day serving as a test to indicate the influence of the prior training on Positions 1 and 2 on performance on Positions 3–8.

Group 2. Group 2 was trained in a relatively homogeneous visual environment

(the dome) and transferred to a very different visual environment (the square room) to determine generalization, if any, from a homogeneous to a heterogeneous environment. The Ss were given seven days of training on Positions 1–8 in the dome (see Fig. 2) with lighting the same as for Group C during transfer-training in the dome. On the eighth day the maze was rotated 45° and the Ss were tested on Positions 1–8. The following day the 7.5-w. bulb was given a coat of black paint, and the bulb and cardboard reflector-disk were encased in two thicknesses of dark cotton cloth to reduce illumination to a point which required approximately fifteen minutes

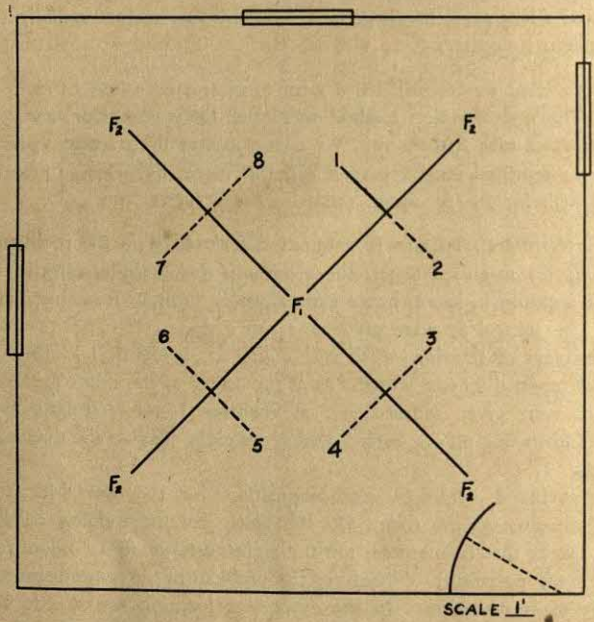


FIG. 3. FLOOR PLAN OF SQUARE ROOM SHOWING DIAGONAL MAZE POSITIONS

for the experimenters to become sufficiently dark-adapted to follow the movement of the S along the maze. Training began for seven days on Positions 1–8 in the square room the following day.

Group 3. To confirm or deny place-learning, Group 3 was trained on four positions in space and then shifted to four new positions. This group was designed to test further whether place-learning can occur or if response was made to relative spatial locations of objects. The maze was in the identical positions it occupied in training Group C and Group 1 and in transfer-training of Group 2; however, the food dish was on the wall-end of the maze (F_2), thereby requiring the Ss to learn the location of four places rather than one. After seven days of training, and upon the eighth day, the maze was rotated 45° and each S was given one run on each of the eight positions.

All training and testing was between 9 A.M. and 4 P.M.

RESULTS

In testing the hypothesis that place-learning is the learning of a definite location in space, a control group of 17 *Ss* was used. By training the *Ss* to go to a definite place by a fixed sequence of turns, RLRLRLRL, the *Ss* could learn the 'place' by either place-disposition or by response-disposition.

Group C. Fig. 4 shows that Group C learned the problem rather rapidly. Perfect runs for the entire group were reached on the seventh day of training, and, as the figure shows, the group had almost reached this level on the fifth day of training.

A series of tests was then given to determine the factors involved in learning the problem. Table I shows that a rotation of maze Positions 1–8 through 45° resulted in 97% correct runs. When these results are compared with those of the first day of initial training, we get a t of 3.14. When testing for place- and response-disposition, the *Ss* were run on Positions 1–8 in a randomized sequence. Although the turning sequence was quite different from the turns in training, only nine errors occurred. A comparison of these results with the initial day of training yields a t of 2.68.

A new maze of identical construction was used the following day as a

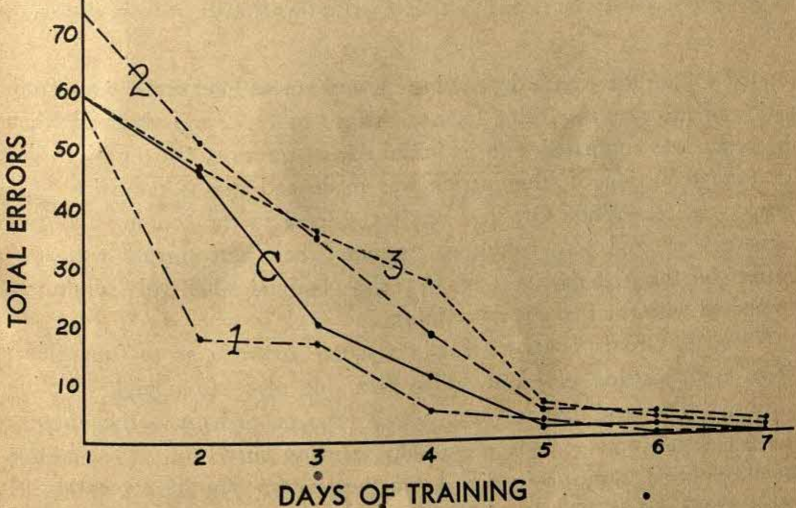


FIG. 4. LEARNING CURVES FOR FIRST TRAINING PERIOD

test of the use of intra-maze cues, and the *Ss* ran the maze without error (see Table I). The *Ss* were then shifted to the dome for additional train-

ing, and there ran the maze with 91% correct choices on the first day (see Table I).

Group 1. Group 1 was trained on Positions 1 and 2 only for seven days. Fig. 4 shows that all *S*s ran without error on the sixth and seventh days. When the group was shifted to Positions 1–8 on the eighth day, 74% correct runs resulted. When compared with their first day of training, a *t* of 3.26 was found, as shown in Table I.

Group 2. Group 2 reached 99% correct runs on the seventh day of training, while running in a homogeneous dome. The maze was then

TABLE I

PERCENTAGE OF CORRECT RUNS UNDER TEST-CONDITIONS

Comparison of test-performance with initial day of training is shown by t.

Group	Testing conditions	% correct	t*
C	Rotation of maze 45°	97	3.14
C	Randomized maze positions	93	2.68
C	New maze	100	11.25
C	Transfer to dome	91	2.27
1	Transfer to Positions 1–8	74	3.26
2	Maze rotated 45°	100	13.72
2	Light intensity severely reduced	54	1.51
2	Transfer to square room	76	5.71
3	Rotation to four new places	96	8.42

* A t of 2.60 is significant at the 1-% level; of 1.96 at the 5-% level; and 1.55 at the 10-% level.

rotated 45° on the eighth day, and the group ran without error. An examination of this result in Table I shows that a *t* of 13.72 was obtained when the group was compared with its initial day of training. On the ninth day a severe reduction of illumination was made and the *S*s ran with 54% correct choices. When Group 2 was transferred to the training room of Groups C, 1, and 3 for additional training, their performance was 76% correct on the first day. A *t* of 5.71 was obtained when this result was compared with the first day of training.

Group 3. Group 3 was trained in a square room to go to four places from eight starting positions rather than one place from eight starting positions. The food dish was near a wall (F_2) on each run rather than in the geometric center of the room. The learning curve for the seven-day training period is shown in Fig. 4. On the seventh day the percentage of correct runs was 99. On the eighth day the maze was rotated 45° (see Fig. 3) with the result that the *S*s were running to four new places. Table I shows that the *S*s made 96% correct runs, and a comparison with the first day of training yielded a *t* of 8.42.

Comparison of groups. An analysis of variance was made on all groups for errors on Position 1 for the first day of training. Between-variance was found to be 0.101 for the groups, and variance within the groups was found to be 1.202, which resulted in an F-ratio of 0.084. All groups were also compared on the basis of the initial day of training for each group. These comparisons are given in Table II. All groups were compared for differences in learning during the first seven days. The results are given

TABLE II

COMPARISON OF GROUPS FOR FIRST DAY OF TRAINING

Group	Group 1	Group 2	Group 3
C	.663*	2.46	.00
1	—	2.08	.55
2	—	—	1.83

* A t of 2.60 is significant at the 1-% level; of 1.96 at the 5-% level; and 1.65 at the 10-% level.

TABLE III

COMPARISON OF GROUPS FOR TOTAL ERRORS FOR FIRST SEVEN DAYS

Group	Group 1	Group 2	Group 3
C	.456*	.409	.489
1	—	.851	.869
2	—	—	.079

* A t of 2.60 is significant at the 1-% level; of 1.96 at the 5-% level; and 1.65 at the 10 % level.

TABLE IV

COMPARISON OF GROUPS FOR TOTAL ERRORS FOR THE SECOND SEVEN DAYS

Group	Group 1	Group 2
C	.516*	.627
1	—	.184

* A t of 2.60 is significant at the 1-% level; of 1.96 at the 5-% level; at 1.65 at the 10-% level.

in Table III. Groups C, 1, and 2 were given a second period of training under conditions stated in Method. These results are presented in Table IV and Fig. 5.

Test conditions. When Group C, trained in a square room, is compared with the performance of the *S*s in the dome room, a t of 1.57 is attained. A t of 0.879 was obtained when the performance of Group 1, trained on Positions 1 and 2, is compared with performance on Positions 1–8, while a t of 1.49 shows the relationship of over-all training of Group 2 in the dome and in a square room. These ts are all based on total training for each condition for a period of seven days and serve as additional evidence for generalization from one concrete to another concrete situation.

Discussion

That spatial learning consists of learning of a definite location in space was the hypothesis tested by Group C in this study. The learning curve of the group demonstrates that the group did learn the location in space. Although training was given in sequential order on maze Positions 1–8, a test upon randomized positions showed learning of the problem occurred by place-disposition rather than by response-disposition. This is in agree-

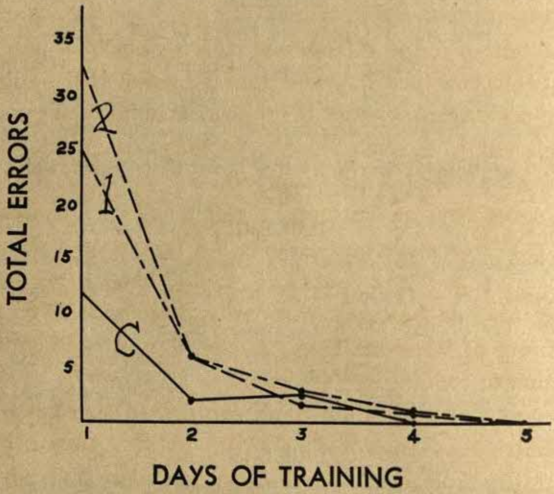


FIG. 5. LEARNING CURVES FOR SECOND TRAINING PERIOD

ment with Ritchie, Aeschliman and Peirce when they state that "simple T-maze learning is based originally on the acquisition of a place disposition."[8]

The high performance of Group C following a rotation of maze-positions by 45° is also indicative of place-learning since the choice-points differed in location, while the food place remained the same. A t of 3.14 indicates a significant difference between this performance and the performance of the group on the first day of training. This difference is significant above the 1-% level; therefore, the prior training facilitated performance at the new choice-points. The performance of this group on a new maze with no errors ruled out the possibility of the use of intra-maze cues in the solution of the problem.

[8] B. F. Ritchie, B. Aeschliman, and P. Peirce, Studies in spatial learning: VIII. Place performance and the acquisition of place disposition, *J. Comp. & Physiol. Psychol.*, 43, 1950, 85.

We cannot, therefore, agree with Woodworth in his definition of place-learning when he states: "By place we mean a concrete situation containing objects in spatial relations,"[9] since the transfer of Group C from a square room containing windows and lighted by sunlight to a circular dome containing no windows and lighted very weakly resulted in a generalization of training. A t of 2.27 between the initial day of training and the first day in the dome is significant above the 5-% level, hence generalization of some type occurred. We can, therefore, say that a definite concrete situation is not required in spatial learning.

If generalization of training from one environment to a quite dissimilar environment can occur, then it would appear feasible that some degree of facilitation of performance would occur from training upon one axis in a certain environment to a given place to another axis in that environment leading to the same place. This assumption was tested in the training of Group 1 for seven days on Positions 1 and 2 and then run on Positions 1–8. When the performance on Positions 1–8 is compared with the initial day on Positions 1 and 2, a t of 3.26 results, hence generalization did occur. Also learning of the eight positions occurred with a greater rapidity than did learning of Positions 1 and 2 only, thus strengthening the idea of generalization of training.

If the foregoing generalization could occur, then a group trained in a relatively homogeneous environment to go to a particular place could possibly generalize from other positions to that place and possibly, also, generalize from the particular environment to a quite dissimilar environment. Following seven days of training, in which *S*s attained 99% correct runs, they were tested on Positions 1–8 after 45° rotation and ran without error.

When this errorless performance was compared with the first day of training, a t of 13.72 resulted, hence generalization did occur from eight new places to the same food place.

Generalization from the dome to a square room occurred despite great differences in room construction and illumination ($t = 5.71$). The conclusion that generalization of learning of a particular location in space in one environment to a particular location in space in a totally new environment is strengthened by the performance of this group.

Since response-disposition was eliminated as the means of learning the spatial problem and since generalizations from one axis to other axes in the same environment transpired, it would appear that spatial learning is

[9] *Op. cit.*, 135.

the learning of finite locations in three-dimensional space; however, this concept was negated by the generalizations of two groups from a definite concrete situation to totally new situations. What appeared to be true spatial learning, then, must result from some other factor. A group trained on four axes to learn four locations in space, and then running to four new locations following a 45° rotation would confirm or deny this postulate by either reverting to chance behavior or would generalize to the four new positions by means of some factor such as a relative location of the maze bifurcation to a wall.

Group 3 learned the four locations but at a somewhat slower rate than Group C learned one location in space on the same axes in the same room. When the maze was so rotated that the food dish now occupied four new positions, maze performance did not drop to chance performance, as it would have had space learning required a finite location in space. The group executed the test with only six errors, a performance of 96% correct runs. A comparison of this performance with that of the first day of training gave a t of 8.42, a significant difference above the 1-% level of confidence. What had appeared earlier to be place-learning of a finite location was not demonstrated.

The test performance of Groups 1 and 2 would appear to indicate strongly that place-learning occurred and that place-learning was dependent on a finite location in space. Since randomized testing of Group C ruled out response-disposition in the problem, this would tend even more to reinforce this apparent conclusion. Furthermore, when two groups were trained in two experimental rooms under different conditions to locate the food position in the center of the rooms from four axes with eight starting-positions, they did so readily, as we have seen. When, however, the two groups were transposed, each to the training environment of the other, a surprising initial performance resulted, particularly surprising since there was a difference in light intensity, wall and ceiling structure, and a totally different visual environment in each room. Inasmuch as the only common factors in the two environments was the relation of the choice-points to the walls, it would seem that what had appeared to be place-learning of a finite location in space was in reality a learning of the food position in relation to a nonspecific wall.

This was readily shown to be the case when a new group, trained to learn four locations in space away from the center of the room, performed with a high degree of efficiency upon rotation of the maze axes 45° so that four new locations, which still led away from the room center, were involved.

The results of the present experiment do not give any clue as to why Gentry, Brown, and Kaplan were unable to confirm the results of Tolman, Ritchie, and Kalish. Gentry, Brown, and Kaplan state: "Since our results for both the animals with training and without training are in part explainable in terms of an artifact of the experimental apparatus (the presence of alley CD), we wonder if their results are not explainable in terms of some artifact of their experimental set-up, and hence if there is any need for them to appeal to the senior author's expectancy theory in its inferential form."[10] Even though Tolman, Ritchie, and Kalish do not indicate the location of the maze with reference to the wall opposite Pathway 6 during training and on the test-trials, it appears extremely doubtful that the position of the walls could account for their results. Tolman and his associates considered the position of the light as the sole visual factor of importance. Ritchie found that *Ss* will orient toward noises coming from a discard cage if one is in the experimental room.[11] In view of this finding by Ritchie, it appears possible that the *Ss* which ran Pathway 6 could have been running to noises coming from a discard cage at the end of Pathway 6, if there was a discard cage at the end of Pathway 6. Since Gentry, Brown and Kaplan returned each *S* to the colony room as it finished its test-trial, their *Ss* could not have performed with reference to a noise.

Summary and Conclusions

This investigation demonstrated that when a group of *Ss* was permitted to learn a spatial location by place-disposition or by response-disposition, learning occurred by place-disposition, a significant result in view of the controversy over the two methods.

Generalization of spatial learning was introduced and studied by several approaches and it was found that learning of a particular location in space from four axes resulted in a facilitation of performance on four new axes leading to the same location in space. Also the learning of a location in space from one axis in a certain environment could be generalized to three new axes in the same environment if the same location in space was maintained. It was further demonstrated that spatial learning in a particular concrete envirnoment could be generalized to a totally new and different environment.

These generalizations led to the conclusion that spatial learning was not the learning of a certain location in space, and the postulation evolved

[10] *Op. cit.*, 321-322.

[11] Ritchie, Studies in spatial learning: III. Two paths to the same location and two paths to two different locations, *J. Exper. Psychol.*, 37, 1947, 25-38.

that generalization of training could occur from four places in a particular environment to four new places in the same environment. This was tested and found to be true. What had appeared to be place-learning was, in reality, response to a relationship between maze bifurcations and non-specific wall surfaces.

The conclusions may be summarized as follows: (1) Spatial learning does not necessarily constitute the learning of a finite concrete location in a particular environment.

(2) Generalization of learning from one or more axes to a given place may occur to other axes leading to the same place.

(3) Learning of location in space in one environment may result in a facilitation of performance at a similar location in a new and different environment.

(4) Learning of four places in a particular environment facilitates the performance at four new places in that environment. This shows that learning occurred by relationships in that environment rather than the learning of locations in space.

(5) While extra-maze visual cues are necessary for space-learning to occur, the extra-maze visual cues need not be specific in nature.

THE REINFORCING EFFECT OF TWO SPOKEN SOUNDS ON THE FREQUENCY OF TWO RESPONSES

By JOEL GREENSPOON, Indiana University

The reinforcing effects of various stimuli presented immediately following a response have been investigated largely with infra-human *S*s. In the context of experiments using the operant conditioning paradigm, the accepted definition of the reinforcing stimulus is a stimulus introduced following a response that increases the probability of occurrence of that response. Despite this research utilizing infra-human *S*s, there has been relatively little effort to identify reinforcing stimuli for human *S*s. Many investigators have conditioned humans, but they have used only a few reinforcing stimuli. Thorndike demonstrated the effectiveness of 'right' and 'wrong' in increasing the frequency of different responses.[1] Hurlock demonstrated that praise and reproof significantly affected performance in the classroom situation.[2] Other investigators have demonstrated that various stimuli would increase the probability of responding.[3] Most of the research involving reinforcing stimuli with human *S*s has, however, been designed to test problems other than the identification of reinforcing stimuli for human *S*s. The primary purpose of this research was to investigate the effect of the introduction and omission of two spoken sounds following a pre-determined response on the frequency of occurrence of that response.

---

* Accepted for publication October 1, 1954. This article is based upon a dissertation submitted to the Department of Psychology of Indiana University in partial fulfillment of the requirements for the Ph.D. degree. This research was supported in part by a grant from the Graduate School of Indiana University. Appreciation is expressed to Dr. C. J. Burke for his aid in the statistical analyses and to Dr. T. A. Sebeok for his aid in the phonetic analysis of the contingent stimuli.

[1] E. L. Thorndike, *The Psychology of Wants, Interests, and Attitudes*, 1935, 10-305.

[2] E. B. Hurlock, The evaluation of certain incentives used in school work, *J. Educ. Psychol.* 16, 1925, 1-49.

[3] A. J. Mitrano, Principles of conditioning in human goal behavior, *Psychol. Monog.*, 51, 1939 (No. 230), 1-55; Louis Long, Conceptual relationships in children: The concept of roundness, *J. Genet. Psychol.*, 57, 1940, 289-315; A. B. Warren and R. H. Brown, Conditioned operant response phenomena in children, *J. Gen. Psychol.*, 28, 1943, 161-207; K. W. Estes, Some effects of reinforcement upon verbal behavior of children, Unpublished doctor's dissertation, University of Minnesota, 1945; P. R. Fuller, Operant conditioning of a vegatative human organism, this JOURNAL, 62, 1949, 587-590; E. J. Hovorka, A study of stimulus-generalization in human operant behavior, Unpublished master's thesis, Indiana University, 1950; B. D. Cohen, H. I. Kalish, J. R. Thurston, and E. Cohen, Experimental manipulation of verbal behavior, *J. Exper. Psychol.*, 47, 1954, 106-110.

PROCEDURE

The experiment was simple in design. *S* was asked to say words, and as he went along, some of the words were followed by a spoken sound from *E*. Conditioning and, later, extinction were both obtained.

The experiment was conducted in a small room, 7 × 7 × 7 ft., with sound-insulated walls, and lighted by one 75-w. ceiling bulb. The room contained a small table and two chairs. *S* sat in one chair placed beside the table and was unable to see *E* who sat behind him in the other chair. A small red light was placed on table where it could be seen by *S*. He could also see a microphone that was attached to a Peirce Wire Recorder. The recorder sat on a small stand out of sight during the experiment, but it was visible to *S* when he entered the experimental room. A stop watch was used to record time.

Seventy-five undergraduate students in elementary psychology and speech classes at Indiana University were randomly assigned to five different groups of 15 *Ss* each. Each *S* was tested individually.

The two contingent stimuli were 'mmm-hmm' and 'huh-uh.'[4] The phonetic construction of these sounds was: ʔm̃ˈhm̃ʼ and ˈhã̄ʔã̄ˑ.
The indication of stress is Pike's.[5] The stress and phonetic pattern represent the norm as each one was not pronounced in exactly the same way each time.

Two responses were defined for use in the experiment. One response included any plural noun. The second response included all verbal responses except plural nouns and is called non-plural responses. The defining characteristic of the plural and non-plural responses was based on common grammatical usage.

The experimental session was 50 min. in length. *S* first entered the experimental room and seated himself. A brief, casual conversation, to acclimate *S* to *E* and to the experimental room, preceded the following instructions:

"What I want you to do is to say all the words that you can think of. Say them individually. Do not use any sentences or phrases. Do not count. Please continue until I say stop. Go ahead."

No additional instructions were given during the remainder of the experimental session. *S* received no information about the correctness of his response or the significance of the contingent stimulus that was introduced.

For Groups I and II the contingent stimulus was introduced following each plural response during the first 25 min. For Group I the contingent stimulus was stimulus was introduced following each non-plural response during the first 25 min. For Group III the contingent stimulus was 'mmm-hmm' and for Group IV it was 'huh-uh.' *Ss* in all groups continued to respond for an additional 25 min. during which the contingent stimulus was omitted. One control group was used, in which 'mmm-hmm' and for Group II it was 'huh-uh.' For Groups III and IV the contingent no contingent stimulus was introduced during the entire 50 min. session. At the end of the 50 min. of responding each S was asked the following questions: (1) What do you think it was all about? (2) Did you notice any change in the

[4] The two stimuli used in this experiment, 'mmm-hmm' and 'huh-uh,' are called contingent stimuli rather than reinforcing stimuli since one of the purposes of the experiment was to determine whether or not these two stimuli were reinforcing stimuli.

[5] K. L. Pike, *American English Intonation,* 1946, 1-135.

kind of words you were saying? (3) What do you think the purpose of the 'mmm-hmm' (or 'huh-uh') was? (4) How long do you think you were saying words?

RESULTS

The first step in the treatment of the data was to eliminate those *S*s who were able to verbalize the relationship between the contingent stimulus and the response which it followed. One *S* in Group I and nine *S*s in Group II reported that they noted the relationship of the contingent stimulus and the response it followed. The elimination of these 10 *S*s reduced to 65 the number for whom the data were further analyzed.

The second step in the analysis was to determine the ordinal position of the first plural response. The control group and the two experimental groups in which the plural response was the measured response were compared. The mean ordinal position of the first plural response of Groups I and II and the control group is presented in Table I. The three values did not differ significantly as an F of 0.0626 with 2 and 32 degrees of freedom was obtained. The groups of *S*s were probably

TABLE I

MEAN ORDINAL POSITION OF FIRST PLURAL RESPONSE OF CONTROL GROUP AND EXPERIMENTAL GROUPS IN WHICH CONTINGENT STIMULUS WAS INTRODUCED FOLLOWING EACH PLURAL RESPONSE

Control Group (No stimulus)	Group I ('Mmm-hmm')	Group II ('Huh-uh'
21.13	23.71	20.83

selected from the same population with respect to the readiness to give the first plural response. No corresponding analysis was made for Groups III and IV.

The total 50 min. of responding was divided into ten 5-min. periods for purpose of the additional analyses. The data of the control group and the experimental groups in which the contingent stimulus was introduced following each plural response are presented first. Both periods, during which the contingent stimulus was introduced and omitted, are included.

A graphic presentation of the mean frequency of all verbal responses, both plural and non-plural, is given in Fig. 1 for the control group and Groups I and II. There is a progressive decline in the mean number of verbal responses for all groups.

The mean number of plural responses for each 5-min. period by Groups I and II and the control group is presented in Table II, with the corresponding standard deviations. The generalized analysis of variance was applied separately to the plural responses during the periods when the contingent stimulus followed each plural response and when it was omitted.[6]

[6] The generalized analysis of variance was developed by C. J. Burke of Indiana University. It can be applied to those cases in which there is correlation between measurements within the various groups. The test is designed to evaluate variance and co-variance simultaneously. The measurements in successive periods of the

The between-group-variance for the periods 1–5 when the contingent stimulus was introduced for the experimental groups was significant beyond the 1-% level of confidence. The between-group-variance was significant between the 5-% and 1-% levels of confidence during the remaining periods when the contingent stimulus was omitted.

The t-test was applied to determine the locus of the differences in the mean number of plural responses between the experimental groups and the control group; every 5-min. period was examined separately. The results of this analysis show that Group I had a significantly greater mean number of plural responses than the control group for the last four

TABLE II

MEAN NUMBER AND STANDARD DEVIATION OF PLURAL RESPONSES FOR SUCCESSIVE 5-MIN. PERIODS FOR CONTROL GROUP AND FOR EXPERIMENTAL GROUPS IN WHICH CONTINGENT STIMULUS WAS INTRODUCED FOLLOWING EACH PLURAL RESPONSE
(Stimulus omitted last five periods of experimental groups)

5-min. periods	Control Group (No stimulus)		Group I ('Mmm-hmm')		Group II ('Huh-uh')	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	15.47	11.60	25.50	22.80	11.33	5.62
2	11.20	9.22	22.07*	13.53	7.17	5.50
3	11.00	6.83	22.43*	16.90	2.83*	2.68
4	10.53	7.74	19.07*	13.19	4.83	3.85
5	8.40	8.93	20.86†	11.36	3.83	3.19
6	8.13	5.77	16.21*	12.18	7.33	6.90
7	8.27	6.50	11.64	9.24	4.83	5.46
8	10.87	10.30	10.50	8.68	3.00	4.40
9	6.67	6.48	11.43	10.11	7.33	6.13
10	8.33	7.94	9.50	7.38	4.83	5.88

* Mean difference between experimental and control groups significant between the 5-% and 1-% level of confidence.

† Mean difference between experimental and control groups significant beyond 1-% level of confidence.

periods in which the contingent stimulus was introduced and for the first period in which the contingent stimulus was omitted. Group II had a significantly smaller mean number of plural responses than the control group in one period in which the contingent stimulus 'huh-uh' was introduced and in none of the periods in which the stimulus was omitted.

Substantially the same analysis was made of the data for Groups III and IV, in which non-plural rather than plural responses were followed by the contingent stimulus. The mean frequency of all responses for suc-

groups used in this experiment are presumably correlated since they are made on the same *Ss*. If the result is significant, then t-test can be used to find the locus of the differences. The results of this analysis are presented in terms of the confidence level from conversion tables developed by Burke. The information necessary to compute the statistic was obtained through personal communication.

cessive 5-min. periods is presented in Fig. 2, and again there is a continuous decline in the rate of responding.

The mean number of non-plural responses with corresponding standard deviations for each 5-min. period for the two experimental groups and the control group is presented in Table III. The generalized analysis of variance indicated that the difference among the groups when the contingent stimulus was introduced was significant beyond the 5-% level of confidence. The difference was not statistically significant during the periods when the contingent stimulus was omitted. The analysis by means of the *t*-test showed none of the mean differences to be statistically significant.

Discussion

The results obtained from the introduction of 'mmm-hmm' were consistent for both of the responses, plural and non-plural. Since, according

TABLE III

Mean Number and Standard Deviation of Non-plural Responses for Successive 5-Min. Periods of Control Group and Experimental Groups in Which Contingent Stimulus Was Introduced Following Each Non-plural Response
(Stimulus omitted last five periods of experimental groups)

5-min. periods	Control Group (No stimulus)		Group III ('Mmm-hmm')		Group IV ('Huh-uh')	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	102.67	34.50	93.93	39.33	95.07	40.50
2	79.40	28.01	84.33	34.99	87.60	37.46
3	75.40	26.49	82.13	36.59	87.00	36.75
4	72.27	26.79	80.80	35.46	84.60	35.38
5	73.60	29.96	79.33	35.46	90.47	41.10
6	70.47	28.11	75.93	34.23	80.20	26.53
7	72.47	26.53	74.73	33.53	76.07	33.47
8	67.73	24.70	70.27	30.67	77.60	32.70
9	70.87	25.03	71.60	33.16	71.20	29.70
10	68.93	28.05	72.87	35.11	68.87	24.47

to the initial definition of the reinforcing stimulus, any stimulus introduced following a response that increases the probability of occurrence of that response is a reinforcing stimulus, we may conclude that 'mmm-hmm' is a reinforcing stimulus. Additional support for this conclusion comes from the results that were obtained when 'mmm-hmm' was omitted. The frequency of plural responses declined to the point where the difference between the mean number of plural responses of the control group and Group I was not statistically significant.

The results obtained from the introduction of 'huh-uh' following the plural response were significantly different from the results obtained when 'huh-uh' was introduced following non-plural responses. The results were

obtained, however, from only 6 *S*s in Group II who did not verbalize the relationship between 'huh-uh' and plural responses. This represents a rather small sample. The apparently diverse results in the case of 'huh-uh' may clarify some of the thinking about reinforcing stimuli. It would

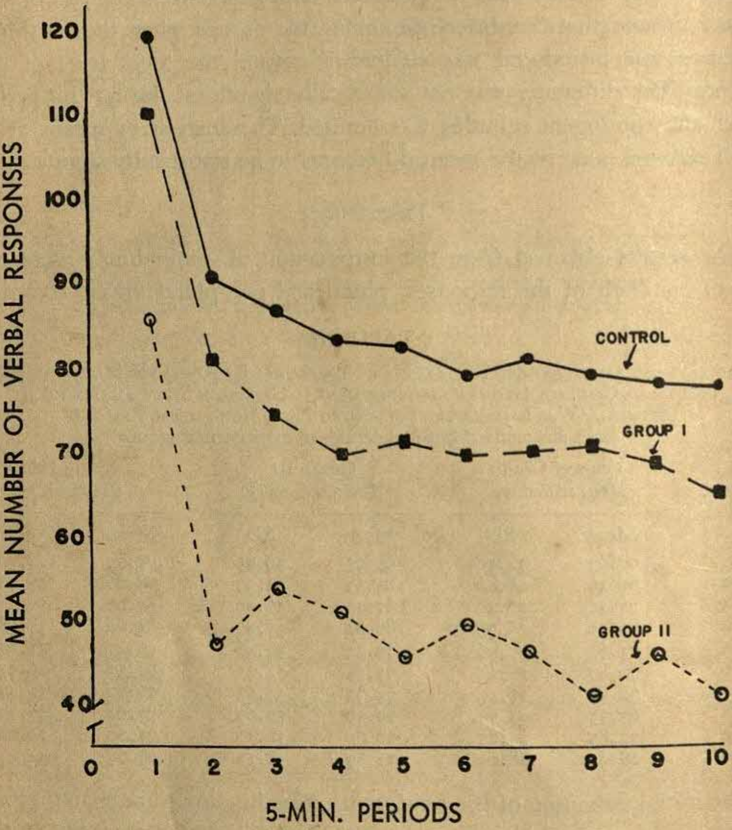


FIG. 1. MEAN NUMBER OF VERBAL RESPONSES FOR SUCCESSIVE 5-MIN. PERIODS OF THE CONTROL GROUP AND THE EXPERIMENTAL GROUPS IN WHICH THE CONTINGENT STIMULUS WAS INTRODUCED FOLLOWING EACH PLURAL RESPONSE
(The contingent stimulus was omitted during the last five periods of the experimental groups.)

appear from these results that one of the factors that may determine whether or not a particular stimulus will be a reinforcing stimulus is the response following which the stimulus is applied. An examination of the two responses used in this experiment reveals some differences. Plural responses are a smaller and more narrowly defined class in that all mem-

bers of the class were plural nouns. The data from the control group indicate that approximately 11% of the verbal responses were plural nouns. Non-plural responses form, therefore, a much larger class. They also presumably differ from the plural responses in being more heterogeneous. All parts of speech other than nouns, and also non-plural nouns, are included. Thus, either the relative size or the heterogeneity of the class, or both, may be factors in determining whether or not a particular stimulus will be a reinforcing stimulus.

It should be noted that there was little tendency for the *S*s to repeat a particular word that had been followed by one of the contingent stimuli.

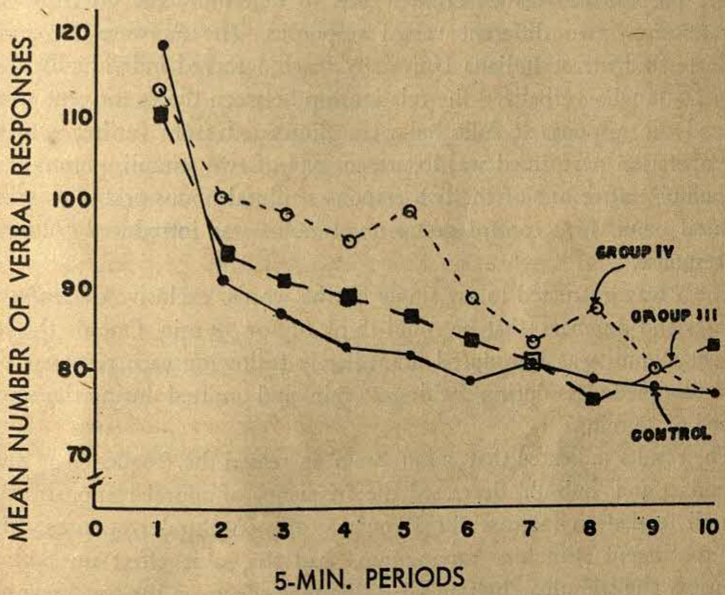


FIG. 2. MEAN NUMBER OF VERBAL RESPONSES FOR SUCCESSIVE 5-MIN. PERIODS OF THE CONTROL GROUP AND THE EXPERIMENTAL GROUPS IN WHICH THE CONTINGENT STIMULUS WAS INTRODUCED FOLLOWING EACH NON-PLURAL RESPONSE (The contingent stimulus was omitted during the last five periods of the experimental groups.)

It was possible for *S* to make responses which differed in many respects but were the same in that they were all plural nouns or were non-plural responses. Thus, the importance of the class is emphasized by this experiment. *E* limits the extent of the class by his use of the reinforcing stimulus; but, the extent of the class may in turn determine whether a stimulus has reinforcing effects.

The small differences in the number of non-plural responses between the control and experimental groups may be a function of the fact that the frequency of non-plural responses of the control group approaches a maximum. Any possible increase in frequency of non-plural responses is restricted when compared to the possible increase in the frequencey of plural responses. This restriction in the size of the possible difference between the control and experimental groups may have reduced the statistical significance of the differences as well.

SUMMARY

The purpose of this experiment was to determine the effect of two operations on two different verbal responses. The *S*s were 75 undergraduate students at Indiana University. Each *S* served individually. Data from 10 *S*s who verbalized the relationship between the contingent stimulus and the response it followed were eliminated from further analyses. The operation performed was to present one of two stimuli, 'mmm-hmm' or 'huh-uh,' after one of the two responses, plural nouns or any word not a plural noun. In a control group no stimulus was introduced following the response.

The *S* was instructed to say singly all the words, exclusive of sentences, phrases and numbers, that he could think of for 50 min. One of the contingent stimuli was introduced immediately following each response of a predetermined class during the first 25 min. and omitted during the second period of 25 min.

The results indicated that 'mmm-hmm' increased the frequency of plural responses and 'huh-uh' decreased the frequency of plural responses. Both stimuli tended to increase the frequency of non-plural responses. Thus, the contingent stimulus, 'mmm-hmm,' had the same effect on both responses. The stimulus, 'huh-uh,' had different effects on the two responses. This differential effect on the two responses suggested that the nature of the response is a determinant of the reinforcing character of the stimulus.

A FURTHER ANALYSIS OF 'LEARNING WITHOUT AWARENESS'*

By EMILY B. PHILBRICK and LEO POSTMAN, University of California

A fundamental assumption of Thorndike's theory of effect is the automatic action of reward in strengthening stimulus-response connections. Thorndike and Rock sought to provide evidence for this hypothesis by demonstrating that "a satisfying after-effect could strengthen the connection which it followed and to which it belonged in cases where the learner did not know what the connection was."[1] Their experiment used the method of free-association, with differential reinforcement for specific classes of response. Sequential or rote responses were reinforced; denotative associations were not. Thorndike and Rock assumed that insight into the correct principle would produce a sudden drop in errors whereas the automatic effects of reward would result in gradual learning. Since the frequency of correct response increased gradually rather than suddenly, Thorndike and Rock regarded their results as evidence for learning without awareness.

Further experiments showed, however, that the slope of the learning curve does not provide a reliable criterion for discriminating between learning with and without awareness. Irwin, Kaufman, Prior, and Weaver repeated the experiment of Thorndike and Rock with one critical modification: *Ss* were informed of the correct principle at different points in the experimental series.[2] Even after explicit enunciation of the principle, the learning curve showed a gradual rise similar to that obtained by Thorndike and Rock. Knowledge of the correct principle does not ensure correct responses. The *Ss* must learn to apply the principle, and the acquisition of the necessary skill may take a gradual course.

A further experimental analysis of the problem posed by Thorndike and Rock was undertaken in a study by Postman and Jarrett.[3] Using the same experimental situation as the earlier investigators, Postman and Jarrett compared the performance of three groups of learners: (*1*) those who never achieved knowledge of the correct principle; (*2*) those who were able to infer and verbalize the correct principle at some time during the learning series, and (*3*) those who were explicitly instructed in the correct principle early in the experiment. Those *Ss* who never

* Accepted for publication October 7, 1954.

[1] E. L. Thorndike and R. T. Rock, Jr., Learning without awareness of what is being learned or intent to learn it, *J. Exper. Psychol.*, 17, 1934, 1.

[2] F. W. Irwin, K. Kaufman, G. Prior and H. B. Weaver, On 'learning without awareness of what is being learned,' *J. Exper. Psychol.*, 17, 1934, 823-827.

[3] Leo Postman and R. F. Jarrett, An experimental analysis of 'learning without awareness,' this JOURNAL, 65, 1952, 244-255.

achieved knowledge of the principle showed only a small and insignificant amount of improvement, although the trend in performance indicated that some learning was slowly taking place. The *Ss* who were eventually able to verbalize the principle showed gradual but significant improvement prior to the point of verbalization; after verbalization of the principle, there was a sharp increase in the number of correct responses, followed by a further period of gradual improvement. Those *Ss* who were informed of the correct principle showed a gradual rise in the number of correct responses, in agreement with previous findings.[4] The results of this study confirm the conclusion that the shape of the learning curve bears no determinate relationship to the presence or absence of awareness. The findings also suggest (1) that *Ss* may gradually learn to respond in accordance with a general principle before explicit verbalization of the principle is possible, and (2) that statement of the principle may greatly facilitate such learning.

Clearly, the question of whether or not learning without awareness occurs in human *Ss* does not have an all-or-none answer. Awareness of the rules governing correct responses may be one important condition of improvement but apparently it is not a necessary one. The importance of the rôle played by awareness may well depend on the complexity of the principle and on the readiness with which the general rule can be applied to specific instances. There is evidence in some recent studies that relatively simple verbal habits, *e.g.* the preferential use of certain classes of words or pronouns, may be effectively strengthened by differential reinforcement without awareness on the part of the *Ss*.[5] The variable of complexity of principle merits further investigation in relation to the problem of learning without awareness.

Two questions suggest themselves. First, will learning without awareness occur more readily if the principle of correct response is relatively simple and easily applicable than if it is relatively complex and difficult to apply? Secondly, will awareness of a simple and unequivocal principle lead to a 'sudden' and almost complete elimination of errors, in contrast to the gradual and incomplete improvement reported in earlier studies? The present study presents some evidence bearing on these two questions.

The principle of correct response was both simpler and more readily applicable than that used by Thorndike and Rock and by subsequent investigators. Since the principle presented no difficulties of application, correct and incorrect responses could be differentially reinforced with much greater certainty on the part of *E* than in the case of the free-association technique employed in the earlier studies. Under these conditions, we might expect (1) that significant improvement would occur in the absence of awareness and (2) that verbalization of the principle would be followed rapidly by errorless performance.

---

[4] Irwin, Kaufman, Prior, and Weaver, *op. cit.*, 823-827.

[5] B. D. Cohen, H. I. Kalish, J. R. Thurston, and Edwin Cohen, Experimental manipulation of verbal behavior, *J. Exper. Psychol.*, 47, 1954, 106-110. See also Joel Greenspoon, The effect of two nonverbal stimuli on the frequency of members of two verbal response classes, *Amer. Psychologist*, 9, 1954, 384 (Abstract).

METHOD

Experimental procedure. The experimental method was closely modeled after a procedure frequently used by Thorndike in investigations of the law of effect.[6] The *Ss* are presented with a series of words and are required to respond to each word with a number. Some responses are called 'right,' the other responses are called 'wrong.' The stimulus-words used in this study varied in length from two to ten letters, and *S* had to respond to each word with a number from one to nine. The response was called 'right' if the number was equal to the number of letters in the stimulus-word *minus* one, *e.g.* test—3, telephone—8, and so forth. All other responses were called 'wrong.'

There were 216 different stimulus-words. The words were arranged in 24 blocks of 9 each, with one word of each length placed at random in each block. The order of blocks was systematically rotated in order to eliminate any possible effects due to the serial position of specific stimuli. Each of the 24 blocks was used as a starting block for two *Ss*, so that the total experimental sample consisted of 48 *Ss*.

Presentation of the stimulus-series was terminated if *S* reached a criterion of two successive errorless blocks before the end of the series. The total series of 24 blocks was always presented to *Ss* who did not reach this criterion.

At the end of every block in which *Ss* gave evidence of responding at a significantly better than chance level, they were interrupted and asked to state the principle on which they were basing their responses. These inquiries continued until *S* had stated the correct principle. The criterion of deviation from chance was four or more correct responses in a single block. With nine numbers to choose from and nine stimulus-words, the number of correct responses expected by chance is 1.0, with a standard error of 0.94; 4 correct responses are 3.13 standard errors beyond chance. Even if the estimate of chance is distorted by number-guessing habits,[7] the criterion is so stringent as to be considered adequate. Once *S* had stated the correct principle, no further inquiries were made, but presentation of the stimulus-series was continued until the criterion of two errorless blocks was reached or the total series had been presented.[8]

Instructions. At the beginning of the experiment, each *S* was given the following instructions:

I have prepared a list of pairs of words and numbers. I am interested in seeing how many of the paired numbers you can guess correctly. I will show you the words, and you guess the numbers that go with them. I have the numbers written on the paper on which I shall record your guesses. I have used numbers from one through nine, so you can choose from one through nine for each word. With this large selection of numbers you will probably not make many correct guesses. There are 216 different words, and I will go through the list just once. Try to give your responses as quickly as you can. I will say, 'right' or 'wrong' after each response.

The presentation of the stimulus-list was then begun. The rate at which the words

---

[6] Thorndike, *An Experimental Study of Rewards*, 1933, 1-72.

[7] For a discussion of number-guessing habits, see W. O. Jenkins and Leta Cunningham, The guessing-sequence hypothesis, the 'spread of effect,' and number-guessing habits, *J. Exper. Psychol.*, 39, 1949, 158-168.

[8] There was one case, in which *S* stated the correct principle on the 23rd trial, in which the final criterion was one rather than two errorless blocks.

were presented depended on the speed of *S*'s responses. The average rate of presentation was approximately 4 sec.

Subjects. The *S*s were 48 adults, between the ages of 25 and 55 yr., who were known to *E* through community connections and volunteered to take part in the experiment. They did not know the purpose of the investigation.

RESULTS

There were 20 *S*s who were able to verbalize the correct principle sometime during the learning series; the other 28 *S*s were never able to state the rule governing the correct responses. The results for the successful and unsuccessful groups will be presented separately.

Principle verbalized. The *S*s who were able to verbalize the correct principle did so at different points in the learning series. The point at which the principle was successfully verbalized varied from Block 2 to Block 23. The average number of blocks preceding correct statement of the principle was 11.30. Thirteen of the 20 *S*s reached the criterion of 4 or more correct responses per block well before they were able to verbalize the correct principle. For the entire group, there was an average of 4.25 blocks between the first time the criterion was reached and correct statement of the principle. Fig. 1 shows the average number of correct responses as a function of distance from the point of correct verbalization. For each block of a given *S*, the distance from the point of verbalization of the principle was determined, and the number of correct responses was averaged for all blocks occupying a given serial position. Since *S*s verbalized the principle at different points in the series, successive points are based on varying numbers of *S*s. There were not enough cases to justify computation of averages beyond the points 15 blocks before, and four blocks after, statement of the principle. With a changing number of *S*s, there is, of course, considerable fluctuation from block to block. Nevertheless, there is a steady rise in the number of correct responses well before the statement of the correct principle. It is also notable that the curve shows positive acceleration as the point of verbalization is approached. When only blocks preceding statement of the principle, *i.e.* Blocks -15 to -1, are considered, the slope of the regression line is significantly different from zero ($F = 32.44$, $p < .01$). Significant learning takes place prior to verbalization of the principle.[9]

[9] During the pre-verbalization period the average number of correct responses per block lies below the criterion of 4. The number of correct responses given during the total series of blocks is, however, well in excess of the number expected by chance ($CR = 20.96$, $p < .01$). This fact, as well as the significance of the slope of the regression line, supports the conclusion that there was a reliable amount of learning.

There is a steep increase in the number of correct responses on the block at the end of which the principle is correctly stated. The difference between Block -1 (3.80 correct responses) and Block 0 (7.05 correct responses) is highly significant ($t = 5.80$, $p < .01$). Statement of the principle is not, however, immediately followed by errorless performance.[10] These findings parallel those obtained by Postman and Jarrett with

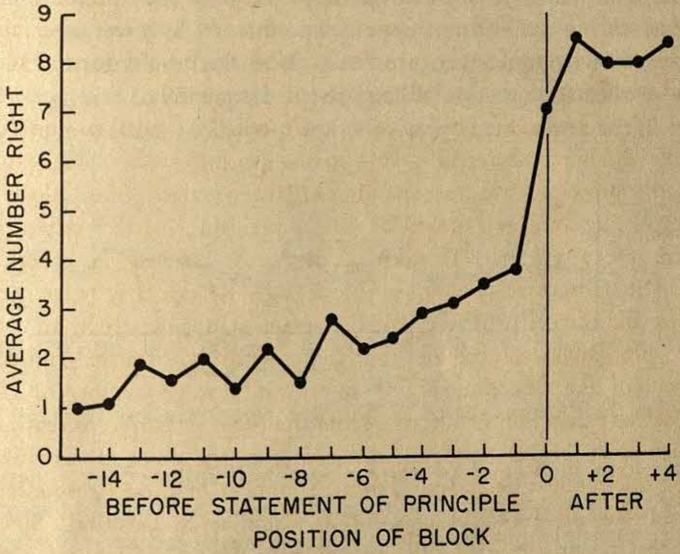


FIG. 1. PERFORMANCE OF *Ss* WHO CORRECTLY REPORTED THE PRINCIPLE
Average number of correct responses on blocks of stimulus-words at varying distances from the point at which the principle was first reported correctly.

a free-association procedure. The level of final improvement is, however, considerably higher in the present investigation than it was in the earlier study. The *Ss* in the study of Postman and Jarrett rarely exceeded a level of 50% correct responses. In this study, all *Ss* who were able to verbalize the principle eventually reached the criterion of errorless performance. The speed and degree of improvement resulting from awareness of the principle depend on the readiness with which the principle can be applied to specific instances.

Principle not verbalized. Fig. 2 presents the learning curve for the 28 *Ss* who did not succeed in verbalizing the correct principle. Although there is considerable fluctuation from block to block, the number of correct

[10] Errors following statement of the principle must be ascribed to failures to discriminate the number of letters in the stimulus-word.

responses shows a steady upward trend. The slope of the regression line is significantly different from zero at the 1-% level of confidence ($F = 7.95$).[11] In the previous study, with the free-association technique, the comparable group of *S*s did not show significant improvement.

The average curve presented in Fig. 2 does not fully reveal the amount of learning in the unsuccessful group. Examination of the individual protocols shows that of the 28 unsuccessful *S*s, only ten never reached the criterion of 4 or more correct responses per block which was used to define significant deviation from chance. Seven *S*s reached this criterion once, and 11 *S*s reached the criterion at least twice, frequently at widely separated points in the series. Many protocols show periods of gradual improvement

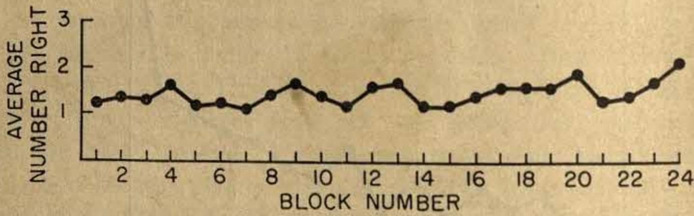


FIG. 2. PERFORMANCE OF *S*S WHO DID NOT STATE THE PRINCIPLE

leading to a high level of correct responses, followed by a period of decline. In such cases, *S*s were gradually learning to respond in accordance with the principle, but the full development of the habit was prevented by interference from incorrect systematic response-tendencies. Examination of *S*s' incorrect hyotheses supports this interpretation. In several cases, hypotheses at least partly related to the length of the words were abandoned in favor of wholly incorrect principles based on such factors as meaning, hedonic tone of the words, and so forth.

Periods of systematic improvement occurred over different parts of the series for different *S*s. Fig. 3 shows the cumulative number of blocks meeting the criterion of 4 or more correct responses as a function of serial position. The curve rises steadily over the entire series of blocks, *i.e.* new cases are added at almost every position. Since periods of improvement were frequently followed by systematic periods of decline, the partial successes achieved by different *S*s canceled each other out to some extent and were not adequately reflected in the overall average of the group. The curve of Fig. 3 does, however, show clear positive acceleration as the

[11] Again the total number of correct responses given over the series of blocks is well beyond chance expectation ($CR = 13.06$, $p < .01$).

end of the series is approached. This increased coincidence in periods of improvement of different *Ss* is reflected in the terminal rise of the average learning curve.

Comparison of groups. Both the group which eventually succeeded in verbalizing the principle and the unsuccessful group showed significant learning without awareness, but the eventually successful group improved at a faster rate *prior to verbalization* of the principle than did the unsuccessful group over the total series of learning trials. On the block preceding statement of the principle, the successful group averaged 3.80 correct responses, whereas the mean number of correct responses on the

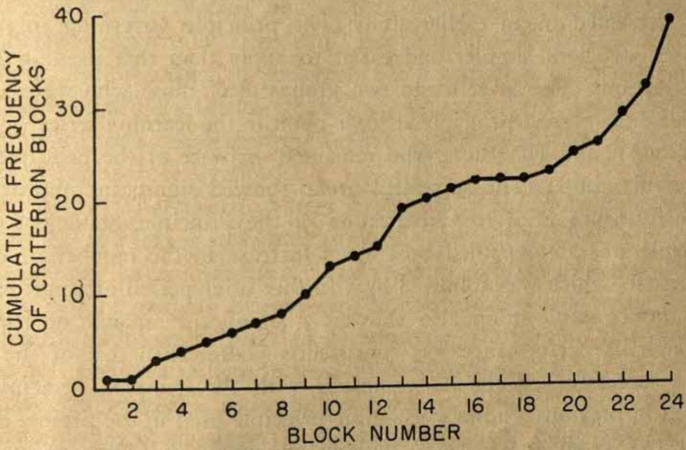


FIG. 3. CUMULATIVE FREQUENCY OF CRITERION BLOCKS FOR THE UNSUCCESSFUL GROUP

last block of the unsuccessful group was only 2.07. The slope of the regression line fitted to the pre-verbalization blocks of the successful group is significantly steeper than that of the regression line fitted to the learning curve of the unsuccessful group ($CR = 5.09$, $p < .01$). The early divergence of the two groups indicates that realization of the principle is related to the degree of initial improvement. Verbalization of the principle signals the fact that a certain critical amount of learning has taken place; thereafter improvement in performance proceeds at an accelerated rate (Fig. 1). 'Insight' into the principle cannot be considered merely as a condition of correct performance; it is at the same time a result of past improvement and a condition of further improvement.

The difference between our successful and unsuccessful *Ss* appears to

reduce to the operation of individual differences. The initial improvement of the unsuccessful group is slower than that of the successful group. At the end of the learning series, the unsuccessful group has not achieved the minimal level of improvement required for verbalization of the principle. If the stimulus-series had been extended, it is highly probable that the unsuccessful *S*s would have shown the same pattern of further improvement as did the successful ones. Awareness develops as the degree of learning increases.

SUMMARY

This paper extends an experimental analysis of 'learning without awareness' presented in an earlier study. The principle governing correct responses was both simpler and easier to apply than that used in earlier investigations. The *S*s fell into two groups: (a) those who were able to verbalize the correct principle at some point in the learning series (successful group), and (b) those who remained unaware of the principle (unsuccessful group). The successful group showed significant improvement in performance prior to verbalization of the principle. Statement of the principle was accompanied by a steep increase in the number of correct responsese, which was followed by a further brief period of improvement. The unsuccessful group also showed a significant amount of learning, although its performance was consistently poorer than that of the eventually successful group. As compared with the complex principle employed in previous studies, the simple principle used in the present experiment (a) facilitated learning without awareness and (b) enhanced the degree of improvement resulting from awareness.

THE INFLUENCE OF 'SET FOR SPEED' ON 'LEARNING WITHOUT AWARENESS' *

By Robert L. Weiss, University of Buffalo

Experiments on the Law of Effect have been beset with methodological difficulties. Thorndike attempted to show that the mechanical action of satisfiers, or rewards, could strengthen verbal response-tendencies without *S*'s awareness of what was being learned.[1] The assumption was made that awareness would be reflected by a sudden improvement in performance, as might be expected in 'insightful' learning. In Thorndike's original experiment, the *S*s learned to make rewarded responses gradually, and it was therefore concluded that they remained unaware of the principle on the basis of which their responses were rewarded. This conclusion has become the issue of further investigations.[2]

In the typical Thorndikian learning situation, a list of stimulus-words is read to the *S*s who are instructed to respond to each word with the very first association. Irwin, Kaufman, Prior, and Weaver have demonstrated that there may be gradual improvement in performance even when the *S*s are taught the basis upon which their responses are reinforced.[3] Since awareness did not produce a *sudden* increase in the number of correct responses, Thorndike's experimental conclusion is open to question. The results of the study by Irwin et al. suggest the hypothesis that the instruction to give the very first association establishes a 'set' for speed which is incompatible with the application of the principle of reinforcement in that it gives *S* no opportunity to translate awareness of the principle into correct performance. The present study was designed to demonstrate the extent to which a 'set for speed' interferes with *S*'s application of the principle of reinforcement. The experimental question was: Under what conditions and in what situations is verbalization of the correct principle of reinforcement essential to improvement.

* Accepted for publication August 16, 1954. The author is grateful to Dr. B. R. Bugelski for many helpful suggestions.

[1] E. L. Thorndike and R. T. Rock, Jr., Learning without awareness of what is being learned or intent to learn it, *J. Exper. Psychol.*, 17, 1934, 1-19.

[2] For a summary of the research that has been done to date, see Leo Postman and R. F. Jarrett, An experimental analysis of 'learning without awareness,' this Journal, 65, 1952, 244-255.

[3] F. W. Irwin, K. Kaufman, G. Prior, and H. B. Weaver, On 'learning without awareness of what is being learned,' *J. Exper. Psychol.*, 17, 1934, 823-827.

In the design of the present investigation, two important methodological errors were avoided: (1) The principle of reinforcement, as used in the previous studies, has been shown to be ambiguous by Postman and Jarrett, who demonstrated that different *E*s obtain significantly different results in applying the original principle of *rote* versus *sequential* associations.[4] A sequential association has been defined as one that completes the meaning of the stimulus-word, *e.g.* across–*the street,* throw–*the ball.* (2) The extent to which the stimulus-words elicit correct responses in the absence of reinforcement has not been indicated in previous experiments. Both criticisms, if not met, limit the extent to which the results can be generalized.

METHOD

Learning material. Twelve blocks of stimulus-words were constructed, with 20 words per block. The 240 words used were selected from *The Teacher's Word Book*[5] on the basis of two criteria; (*1*) all words were verbs, and (*2*) none was beyond the fourth-grade level of difficulty. The 240 words were randomly assigned to the 12 blocks. To minimize any systematic effects of order, the blocks were presented in different orders to different *S*s following a 12 × 12 Latin-square constructed by the method of Bugelski.[6]

Principle of reinforcement. Any response which fell into the verbal response-class 'living things' was reinforced by *E*'s saying 'right,' while all other responses were followed by 'wrong.' *Positively* reinforced responses denoted humans, animals, plants, and the respective parts thereof; names of human characters also were positively reinforced. All pronouns, *e.g. you, self, she, him,* were *negatively* reinforced. (The pronoun *someone* was positively reinforced.)

Experimental groups. Three experimental groups and one control group were used. The experimental groups differed with respect to a *timed-not timed* dimension (whether or not they operated under a set for speed) and an *informed-uninformed* dimension (whether or not they were taught the principle of reinforcement).

The following instructions were read to each *S* in the *informed-timed* group:

I am going to read you a list of words. After each word I say, I want you to say the very first word or phrase that comes to your mind. Some of the words you say will fall into a category of words I have decided to call 'right.' If you say such a word, that is, if your word falls into that category, I will say *right.* If on the other hand the word you say does not fall into my 'right' category, I will say *wrong.* Now, remember, I want you to say the very first word or phrase you think of. It must be the first, so work quickly.

The first block of 20 words was read to each *S*. After the first block was completed, *E* read the following instructions:

Now I am going to tell you on what basis I am calling your responses 'right' or 'wrong.' Any word or phrase that you say which falls into the class of 'living

[4] Postman and Jarrett, *op. cit.*, 251 ff.

[5] E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words,* 1944, *passim.*

[6] B. R. Bugelski, A note on Grant's discussion of the Latin square principle in the design and analysis of psychological experiments, *Psychol. Bull.*, 46, 1949, 49-50.

things' is called right. For example, if I said *run,* and you said *lions,* that would be right. If I said *black,* and you said *green,* that would be wrong. Thus, anything you say that is a living thing I call right. Now we will continue as before, but I still want you to give me the very first word or phrase that comes to mind. It must be the first, so work quickly.

The treatment of the *uninformed-timed* group differed from the former in that the *Ss* were not told the correct principle. The first set of instructions given above was read, and then each *S* was merely asked to state the principle of reinforcement after each of the 12 blocks. *E* could thus determine at what point in the experiment the *Ss* became aware of the correct principle. Whenever this procedure was used *Ss* were told that *E* would not confirm their hypotheses.

The *Ss* of the *uninformed-not timed* group were read instructions similar to those for the informed-timed group except that there was no reference to the response being the first association. On the contrary, *Ss* were told that their response did *not* have to be the first thing they thought of. These *Ss* also were asked to state the principle after each block. The performance of a control group (*uninformed-timed-unreinforced*) provided a measure of the frequency of 'living-thing' responses.

Reliability of scoring. Seven randomly selected protocols were rescored by another scorer.[7] A total of 1680 responses were thus rescored, providing a check on the ambiguity of the principle of reinforcement.

Subjects. There were 12 *Ss* in each of the experimental groups and 47 *Ss* in the control group. The *Ss* were drawn from classes in introductory psychology.

RESULTS

Group performance. To what extent do the four groups differ in performance, as measured by the average number of correct responses per block of stimulus-words? The data are presented in Fig. 1. Both the rate of improvement and the total number of correct responses is greatest for the group *not* operating under the influence of a set for speed. Since the control group does not show a consistent change in performance from the first to the last block, it can be concluded that the amount of improvement in the performance of the experimental groups is a function of the reinforcement.

An analysis of variance performed on the data of the three experimental groups indicated: (1) significant differences—1-% level—among the three groups in total number of responses correct; (2) significant differences among the blocks—1-% level; and (3) a significant interaction term, indicating that the untimed group improved at a faster rate than the other groups—also significant at the 1-% level of confidence.

To what extent does the increase in the average number of correct responses for any single group represent a reliable degree of improvement? The difference between the average number of correct responses for the

[7] The author is indebted to Mrs. Susan R. Meyer who did the rescoring.

first and last third of the blocks, *i.e.* Blocks 1–4, and 9–12, will serve as the measure of within-group improvement. For the *uninformed-timed* and the *uninformed-not timed* groups, the respective ts are 3.5 and 6.2, significant beyond the 1-% level. The means do not differ significantly ($t = 0.60$) for the *informed-timed group,* suggesting no consistent im-

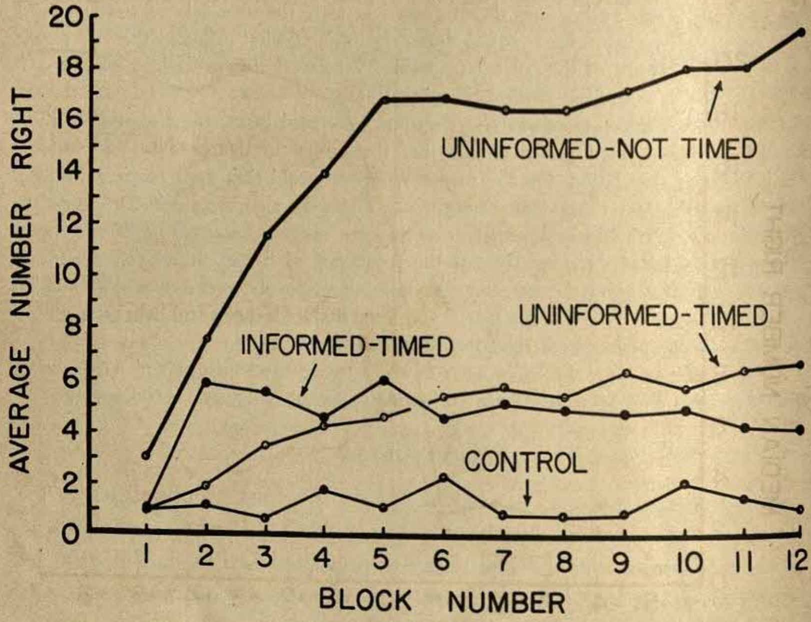


FIG. 1. PERFORMANCE OF THE EXPERIMENTAL AND CONTROL GROUPS

provement after the second block of stimulus-words. Immediately after information about the correct principle was given (Block 2), the average number of correct responses increased to a level significantly greater than that of Block 1 (the t for the mean-difference being 3.5, 1-% level.)

Conditions essential to improvement. To demonstrate the effects of awareness on performance, an analysis was made of the data for the uninformed groups as follows: The block of stimulus-words after which verbalization of the correct principle occurred was taken as the reference point. Median performance on the blocks prior to and following this point was plotted as a function of distance from the point. The curves are shown in Fig. 2. They demonstrate that discovery of the principle of reinforcement had a differential effect on performance under the two conditions of set. For the *uninformed-timed* group (presence of set for speed), the

amount of improvement from the block immediately preceding to that immediately following awareness was three correct responses (*cf.* points -1 and $+1$ in Fig. 2). For the *uninformed-not timed* group (absence of set for speed), the corresponding increase was 15.5 correct responses. The difference in improvement is significant beyond the 1-% level ($t = 4.20$). Awareness of principle seems, therefore, to be essential to

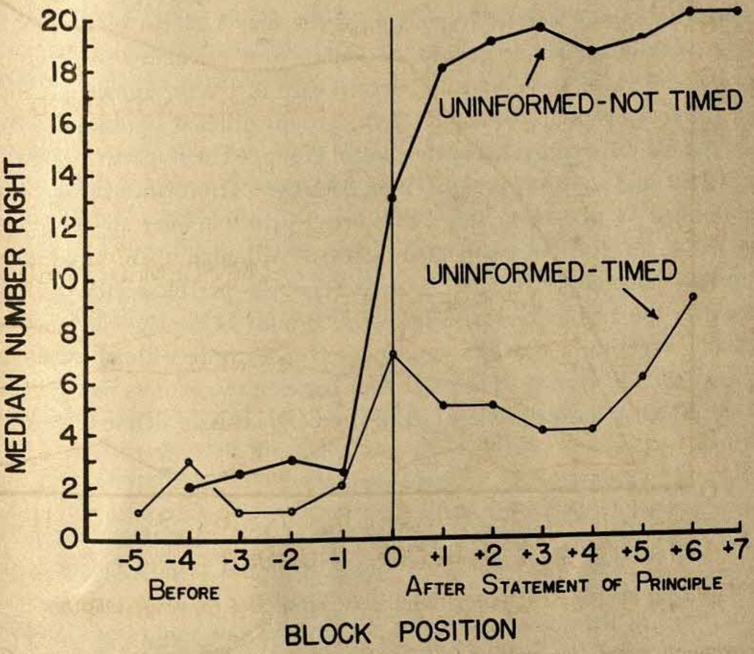


FIG. 2. PERFORMANCE OF THE UNINFORMED GROUPS BEFORE AND AFTER STATEMENT OF THE PRINCIPLE

improvement in the absence of a set for speed.[8] Under these conditions, application of the principle is possible.

Two sources of evidence for 'learning without awareness' may now be evaluated: (1) the performance of the uninformed *Ss* prior to statement of the principle, and (2) the performance of the uninformed *Ss* who never discovered the correct principle.

The performance of the uninformed groups prior to verbalization of the correct principle is not a completely adequate index of learning with-

[8] With the exception of three *Ss* in the *uninformed-timed* group, all *Ss* were able to state the principle at some point in the experiment.

out awareness, for as Postman and Jarrett have correctly observed, *Ss* evolve hypotheses which approximate the correct principle, and their performance reflects partial awareness of it.[9] In the past, such measures have not been compared to control group functioning, and the psychological significance of such performance has accordingly been obscured.

Positions -5 to -1 (Fig. 2) indicate the performance of the uninformed groups prior to statement of the correct principle. Using the median number of correct responses for the first 5 blocks of the control group as the expected frequency, X^2 values were calculated to determine the significance of the difference between each of the experimental groups and the control group.[10] Neither of the groups differed significantly from the expected values provided by the control group. (The respective X^2-values were 1.62 and 7.50, $df = 4.0$.) Nor does the performance of the three uninformed *Ss* who never stated the correct principle offer any support to the notion of learning without awareness. Although their median performance sometimes exceeded 5 correct responses per block, the performance over the 12 blocks was inconsistent. It must be concluded, therefore, that this experiment provides no evidence for learning without awareness.

Ambiguity of scoring. The agreement between two scorers was 99.16%, indicating that the principle of reinforcement, as here defined, is unambiguous.

Discussion

Learning and performance. The results of this experiment show that the manifestation of awareness of the principle of reinforcement is prevented by a set for speed, although the importance of awareness is clearly manifested in the absence of such a set. In previous studies of this problem, performance was taken as the index of learning, but the results of the present experiment demonstrate the influence on performance of a factor which is independent of learning. The performance of the *uninformed-not timed* group, following awareness of the correct principle, suggested insightful learning; the performance of the *uninformed-timed* group might lead us to conclude that learning was gradual; and the performance of the *informed-timed* group might lead us to the anomalous conclusion that learning was first insightful, then followed by a gradual decrement in

---

[9] Postman and Jarrett, *op. cit.*, 249.

[10] Since the *Ss* of the *uninformed-not timed* group inferred the correct principle early in the experiment, there were too few cases to warrant calculation of the median number of correct responses beyond the fourth block prior to statement of the principle. The number of cases decreases with distance from the zero-position in both directions.

efficiency. Can we conclude that there are three types of learning? Certainly not, for we have only three instances of performance affected differently by a non-learning variable.

The problem of awareness. Like Postman and Jarrett, we have employed two criteria of awareness: (*1*) explicit instruction of *S* by *E;* and (*2*) *S*'s ability to verbalize the principle. We also have found that partial awareness influences the performance in an unknown fashion. *S*s may make correct responses for the wrong reason, *i.e.* their hypotheses may be remotely related to the correct principle. The introduction of a control group does not eliminate this problem, for it only provides the average number of correct responses when *S*s do not respond according to one or more hypotheses. It does not provide us with the average number of 'correct' responses when these are determined by incorrect hypotheses. These methodological considerations lead us to conclude that 'learning without awareness' has not been the major issue of previous studies. We agree with Postman and Jarrett that acquisition of verbal responses is not the proper vehicle for demonstrating 'learning without awareness.'

SUMMARY

The present experiment was designed to determine the conditions under which awareness of what was being learned would be essential to improvement. In a Thorndikian verbal-learning situation, performance under three conditions was studied: (1) *S*s were explicitly informed of the principle of reinforcement and given a set for speed, *i.e.* instructions to respond to each stimulus-word with the first association; (2) uninformed *S*s were given the set for speed; (3) uninformed *S*s were not given the set for speed. An uninformed, unreinforced group set for speed served as a control. The results suggest that the utilization of the correct principle is impaired by the set. No conclusive evidence for learning without awareness has been provided by experiments of this type.

A FURTHER STUDY OF PREDICTION-MOTION

By ROBERT M. GOTTSDANKER, *Santa Barbara College*

In the study of *prediction-motion,* *S*s continue to 'track' a target after it has disappeared from view.[1] They are instructed to behave as though the target were an airplane which has flown behind a cloud. Under these conditions, the *S*s show 'smoothed' linear continuation of accelerated patterns, accuracy sufficient to meet typical tracking requirements, and consistent individual differences. In previous experiments, which required continuation of but one rate of movement, the predictive performance of the *S*s was uncorrelated with general tracking skill. The present experiment, which involves several rates of movement, represents a further search for prognostic measures, as well as an attempt to generalize the previous findings.

METHOD

Subjects. The *S*s were 16 students at Santa Barbara College. There were 8 men and 8 women, all between the ages of 18-28 yr.

Apparatus. The apparatus has been described elsewhere.[2] It was a 'tracking box' in which sheets of paper 8½ × 11 in, were driven lengthwise under a ⅛-in. slit in an opaque mask at a constant speed of 22 mm./sec. A pair of adjacent lines 5 mm. apart on each sheet provided a 'target' moving in the slit, which *S* tracked by keeping a pencil between the lines. The pencilled line provided the record of *S*'s responses.

Patterns. Two kinds of pattern were used. The *tracking* pattern involved only ordinary tracking; the target did *not* disappear. It was presented on a series of eight

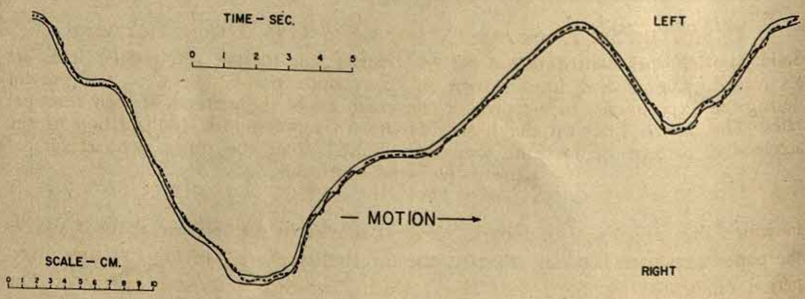


FIG. 1. PORTION OF TRACKING PATTERN AND TYPICAL RECORD

sheets joined end to end, and the total time was 73 sec. A portion of the pattern (two sheets) is shown in Fig. 1.

* Accepted for publication August 16, 1954. Financial assistance for this project was provided by a research grant from the University of California.

[1] R. M. Gottsdanker, The accuracy of prediction motion, *J. Exper. Psychol.*, 43, 1952, 26-36; Prediction-motion with and without vision, this JOURNAL, 65, 1952, 533-543.

[2] Gottsdanker, *op. cit.*, *J. Exper. Psychol.*, 43, 1952, 26-36.

Six *predictive* patterns were used which require *S* to track predictively in the unmarked slit after the termination of the printed pattern. These patterns included three constant rates of motion, 7.5, 11.0, and 20.0 mm./sec.; and three positively accelerated motions, which had the following equations: $s = .484t^2 - .352t$; $s = .871t^2 - .440t$; $s = 1.742t^2 - .440t$; where s is distance in millimeters and t is time in seconds after the target started to move. The six predictive patterns are shown in Fig. 2. It may be seen that there was an initial period of about 1.5 sec.

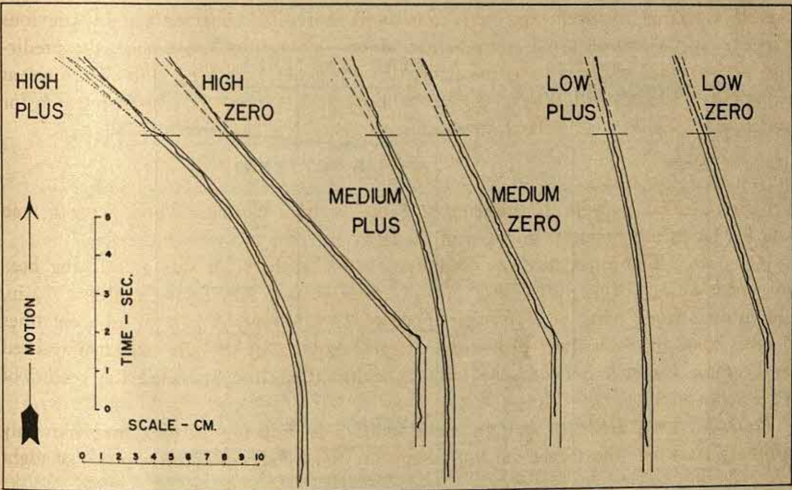


FIG. 2. THE SIX PREDICTIVE PATTERNS AND TYPICAL CONTINUATIONS

Solid parallel lines constitute the pattern tracked; the lighter, less regular lines are *S*'s record. The dashed lines shown in the continuation-period were not present during the experiment; they represent continuations of the patterns at their terminal rates. The dotted lines on the high-plus pattern represent the continuation of the accelerated pattern of motion; they are omitted from the other two accelerated patterns to avoid confusion.

in which the target was stationary. Next, it moved to the left for about 7 sec. as the paper advanced 160 mm. Finally, the target disappeared, giving rise to a 2-sec. period (indicated by the dashed lines) during which *S* tracked predictively. The rates of movement at the instant the targets disappeared were 6.67, 12.23, and 24.90 mm./sec. for the three accelerated patterns. As in the previous studies, the non-accelerated patterns will be referred to as the *zero* patterns and the positively accelerated patterns as the *plus* patterns. In addition, the terms *low, medium,* and *high* will refer to the patterns with respect to terminal rates. Thus, the pattern with the constant rate of movement of 11 mm./sec. is the zero, medium-rate pattern.

Procedure. *S* was given two practice runs on each of the predictive patterns, and then 24 runs in which every pattern occurred four times in random order. After that he undertook two runs on the tracking patterns with the target sheets arranged in a different sequence on each run.

RESULTS

Rates of continuation. The distance covered on every run with the predictive patterns was determined for 2 sec. following disappearance of the pattern (see Fig. 2). Measurement was to the nearest millimeter; the distance divided by 2 gives the rate of continuation in terms of millimeters per second. To permit comparisons between patterns, a percental value was computed, which was accomplished by dividing rate of continuation for a run on a given pattern by the terminal guided rate for that pattern, and then multiplying the quotient by 100. The mean for *S*'s

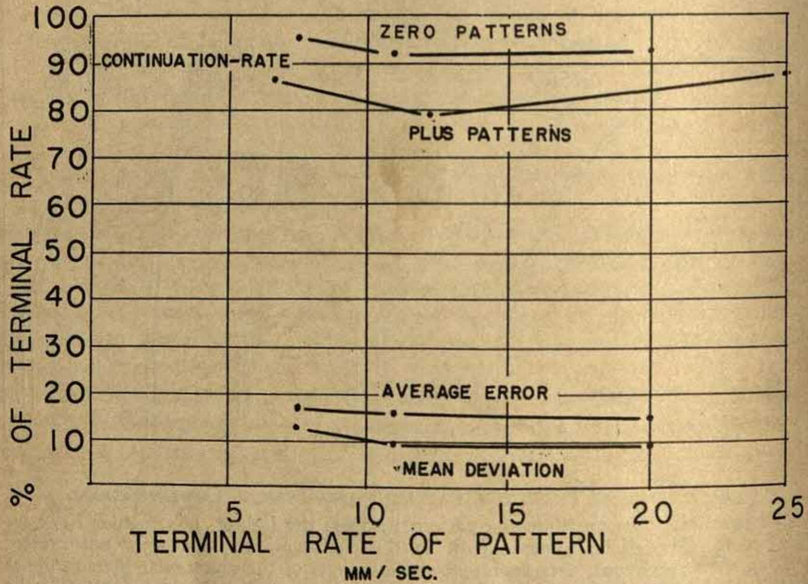


FIG. 3. MEAN VALUES FOR PERCENTAGE-CONTINUATION-RATE, AVERAGE-ERROR, AND MEAN-DEVIATION

Continuation-rate is shown for all six patterns while average-error and mean-deviation are shown for the zero patterns only.

four runs on each pattern will be referred to as *S*'s '*percental rate of continuation*' for the pattern.

Fig. 3 shows the means of the 'percental rates of continuation' for the 16 *Ss* on all six predictive patterns. While the values for the zero pattern are seen to be higher than those for the plus pattern, no clear differences emerge among the low, medium, and high rates. Statistical verification of these indications was provided by an analysis of the variance among the 'percental rates of continuation.' The analysis, which involved a classification into three rates (high, medium, low), two accelerations (zero and plus), and 16 *Ss*, is shown in Table I. It may be noted that the variance due to *Ss* (indicating consistent individual differences) as well as that associated with acceleration was highly significant.

Accuracy. For each of the three zero patterns, *S*'s average error of continuation from the pattern's constant rate was computed for his four runs. This value was converted into a percentage of the pattern-rate to permit comparisons between patterns. *Average-errors* in percentages for the three zero patterns may be compared in Fig. 3. While there is some indication of a downward trend, as the values are 17.2, 16.2, and 15.4% for the low, medium, and high rates respectively, the

TABLE I

ANALYSIS OF VARIANCE AMONG RATES OF CONTINUATION

Source	Sum of squares	df	Variance-estimate	Denominator for *F*-test	*F*
Acceleration (*A*)	2,026	1	2,026.0	(A×R)+(A×S)+(A×R×S)	18.17†
Rate (*R*)	545	2	272.5	R×S	1.50
Subjects (*S*)	16,213	15	1,080.9	R×S	5.98†
A×R	241	2	120.5	A×R×S	1.29
A×S	2,214	15	147.6	A×R×S	1.59
R×S	5,423	30	180.8	A×R×S	1.95*
A×R×S	2,788	30	92.9		
Total	29,450	95			

* Significant at the 5-% level. † Significant at the 0.1-% level.

differences do not prove to be statistically significant. The analysis of variance is shown in Table II. Again, the variance among *S*s is highly significant.

The mean deviation among the rates on the four runs for a given pattern was also computed to obtain a measure of accuracy which, unlike average error, is independent of constant error. Values analogous to those described for average error were determined in percentages and a parallel analysis of variance was

TABLE II

ANALYSIS OF VARIANCE FOR AVERAGE ERROR

Source	Sum of squares	df	Variance-estimate	*F*
Rate (*R*)	28	2	14.0	2.89*
Subjects (*S*)	1783	15	118.9	2.94†
R×S	1216	30	40.5	
Total	3027	47		

* Inverted: R×S/R. † Significant at the 1-% level.

performed on the *mean deviations.* Reference to Fig. 3 again reveals a downward slope. In this instance the variation among the three rates is significant at the 5-% level, and the variance among *S*s achieves the same level. The analysis of variance is shown in Table III.

Prediction of skill in tracking. The criterion of skill in tracking was essentially a time-on-target score for the two runs on the tracking pattern. This score was determined by placing a transparent scoring template having lines ruled at ½-cm. intervals (5/11-sec.) over the tracking record and then counting the number of intervals in which *S*'s tracking line did *not* go outside of the 5-mm. area of the target. Since there were 160 intervals on each run, 160 was the best score obtainable

for one run, and 320 for the two runs combined. The means for Runs 1 and 2 were 85.4 and 93.9 respectively; the *SD*s were 20.9 and 24.1. The two runs, when combined to serve as the criterion, had a mean of 178.1 and an *SD* of 39.5. The coefficient of correlation between the two trials was 0.51, providing a reliability estimate of 0.68 for the combined trials by the Spearman-Brown formula.

A variety of forecasting measures were tested. The two most straightforward ones were obtained in the following manner:

(1) For each of the three zero patterns, the absolute difference between *S*'s 'percental rate of continuation' and the required value of 100 was first obtained.

TABLE III

ANALYSIS OF VARIANCE FOR MEAN DEVIATION

Source	Sum of squares	*df*	Variance-estimate	*F*
Rate (*R*)	121	2	60.5	5.31*
Subjects (*S*)	433	15	28.9	2.54*
$R \times S$	342	30	11.4	
Total	896	47		

* Significant at the 5-% level.

The mean of the three determinations was called *S*'s *continuative rate of error.* For the 16 *S*s the mean was 11.8 and the *SD* was 6.0. The correlation with the criterion was + 0.16.

(2) The three mean deviations were averaged to provide the *combined mean deviation.* For the 16 *S*s the mean was 10.7 and the *SD* was 3.0. The correlation with the criterion was + 0.13.

Neither of the two foregoing measures correlated significantly with the criterion, a value of 0.50 being required for significance at the 5-% level. More complicated measures also were tested against the criterion. Since they too failed to predict the criterion, the necessarily detailed account of their derivation and statistical characteristics may well be omitted.[3]

DISCUSSION

Previous findings were substantiated by the results obtained in the present study. First, positively accelerated patterns were indeed continued at a lower percentage of their terminal rates than were zero acceleration patterns. Thus, additional support is given to the hypotheses that foregoing rates are smoothed or integrated in continuation. It may also be noted that there were no exceptions to this finding within the range of rates employed. Another previous finding which was corroborated was that of consistent individual differences. There was considerable similarity among the three zero patterns in terms of the 'percental rate of continuation' and likewise among the three plus patterns. The evidence, of course, is insufficient to allow the statement that Weber's Law fits this kind of perceptual motor behavior.

The accuracy with which constant rates were continued was found to depend upon

[3] Several measures were based upon the continuations on the plus patterns, and the differences between these continuations and those on the zero patterns. Another approach was to fit a polynomial to the plot representing the relation of *S*'s continuation-distances on the zero patterns to the required distances. A variety of forecasting measures were derived from the parameters of these functions.

the rate which was being continued. There was a trend—albeit a statistically insignificant one—for the average error to decrease as rate increased. This trend did achieve statistical significance when the measure of accuracy used was the mean-deviation. The over-all average-error (16%) was very close to values found in two previous studies (13% and 14%). Whether the slight increase should be attributed to the present use of a right-to-left motion or to a sampling variation cannot be answered from the data at hand. The increase is not sufficiently large to warrant revision of the previous conclusion that predictive motion has enough accuracy to be considered a mechanism which operates in tracking behavior.

As before, the attempt to forecast skill in tracking from measures of continuative reactions was unsuccessful. It may be that prediction-motion has not yet been utilized in the proper way, although a variety of treatments has been employed. Two new approaches to the rôle of prediction-motion should be considered. One is that while the reactions found in prediction-motion do comprise some of the mechanisms underlying tracking behavior they, at the same time, do not provide an important basis for individual differences of skill in ordinary tracking. From this point of view, the rôle of prediction-motion would be similar to that of visual acuity. Other factors, such as speed of response to errors, might be more important in generating individual differences. Both kinds of factor, those having constant and variable effects, ultimately must be incorporated in a satisfactory account of tracking behavior. A second alternative is that the task in the study of prediction-motion, instead of highlighting an essential feature of tracking, is itself new and somewhat restricted. One important difference, for example, is that in prediction-motion *S* does not have the problem of responding to changes of direction as he has in the tracking pattern. From this point of view, little correlation would be expected between the two tasks since it is well known that apparently minor changes in a motor task will drastically affect the ranking of individuals.[4] More might be revealed about tracking behavior by obtaining statistical measures in context rather than by procedures, such as the present one, which attempt rationally to abstract elements from the situation.

SUMMARY

Adult human *S*s were given the task of predictively continuing the tracking of a target which had been made to disappear from view. A pattern of constant rate and a positively accelerated pattern were included within each of three ranges of target-movement. In addition, an ordinary tracking task was given to provide a criterion of skill. The following results were obtained:

(1) Previous results, showing the smoothing of accelerated patterns, relatively high accuracy on patterns of constant rate, and consistent individual differences, were corroborated.

(2) For targets with similar patterns of acceleration, the ratio of continuative rate to terminal rate was fairly constant.

(3) Relative accuracy increased slightly as the rate of the movement of the target increased.

(4) No measure of prediction-motion provided a satisfactory forecast of skill in tracking. The facts argue for a new approach to the problem.

[4] R. H. Seashore, Work and motor performance, in S. S. Stevens (ed.), *Handbook of Experimental Psychology,* 1951, 1345.

THE EFFECTS OF THE ANGULAR RELATIONSHIPS BETWEEN THE OBSERVER AND THE BASE-SURROUND ON RELATIVE DEPTH-DISCRIMINATION

By E. RALPH DUSEK, WARREN H. TEICHNER, and JOHN L. KOBRICK, Quartermaster Research and Development Center, Natick, Massachusetts

A previous study of depth-discrimination by Teichner, Kobrick, and Wehrkamp suggested that the angular relationships between the subject (*S*) and the target's base-surround strongly affect the precision of the settings of equality.[1] Although these experimenters were primarily interested in the distance and type of terrain over which *S* viewed the stimulus-objects, they observed, when the objects were located on rising ground, that *S*'s sensitivity to differences in target-distance was increased. On the basis of this observation they suggested that the contour of the base-surround affects *S*'s ability to discriminate by varying the retinal position of the targets. Although this hypothesis is consistent with their results, they did not have an experimental control. Thus their hypothesis is deductive rather than empirical. In view of the importance of this hypothesis to theories of depth-perception the present study was designed to compare level versus sloping base-surrounds.

Procedure. A Howard-Dolman apparatus, comparable to the one used by Teichner, Kobrick, and Wehrkamp out-of-doors, was constructed. A schematic drawing of it is shown in Fig. 1. The apparatus was painted a flat black except for two white 3-in. square targets. These targets were mounted on hidden parallel tracks on a 48 x 96-in. table whose slope was variable. The horizontal separation between the targets could be varied from 0 to a maximum of 4 in. For this experiment the angle of separation was always 2 min. The required frontal-slopes were obtained by placing blocks of appropriate size under the back legs of the table. The variable target was always on the *S*'s left and *E* controlled its position by means of a tape-chain connection to a 50-in. board placed at the center of the table and extending, as shown in Fig. 1, to the left. A Macbeth Illuminometer and a wooden box 32 in. high were used as accessory apparatus.

The experiment was performed in a large basement which was illuminated by daylight coming through basement windows and by lights placed above the table. Brightness readings of the targets and the base-surround were made daily. The average brightness of the targets was 7.1 ft.-lamberts (apparent ft.-candles) and of the table, taken 12 in. in front, 0.3 ft.-lamberts. The reflection coefficients of the targets and table were approximately 95% and 7%, respectively. A dark 32 x 82-in. cloth centered and hanging 112 in. from the back of the table served as a background.

A factorial design with replication was used, the variables being: (1) viewing distance; (2) frontal-parallel slope of the base-surround of the targets; and (3) height of eye-level. Four enlisted men, inexperienced in depth-experiments, were

* Accepted for publication July 20, 1954.

[1] W. H. Teichner, J. L. Kobrick, and R. F. Wehrkamp, The effects of terrain and observation distance on relaitve depth discrimination, this JOURNAL, 68, 1955, 193-208.

used as *S*s. All had good stereoptic acuity as measured by the Verhoeff Stereopter. The experiment lasted six days with a two-day break between the second and third experimental days. To minimize the effects of practice each *S* was given a preliminary day during which he made 40 target-settings. An attempt was also made to equalize the effects of practice and of changes in morning and afternoon illumination on the several conditions by reversing the sequence of presentation of a second replication of the experiment. Each replication consisted of judgments made with table-slopes of 0%, 3%, and 5%, *i.e.* of 0°, 1.7°, and 2.9°. These variables were

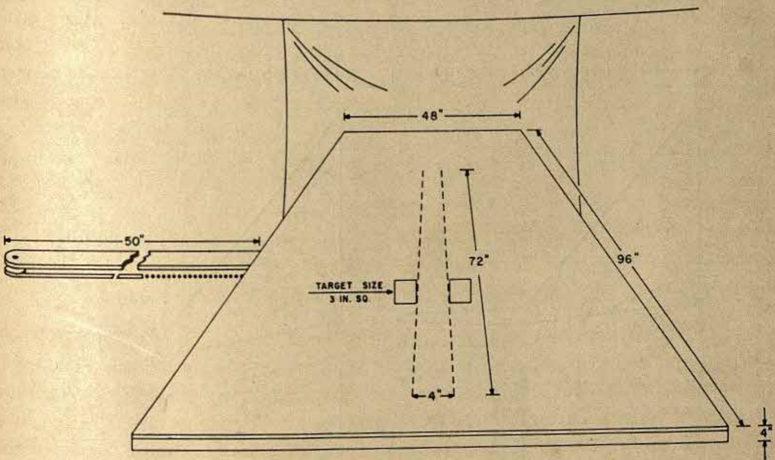


FIG. 1. FRONTAL VIEW OF THE APPARATUS

combined in a factorial design with observation at distances of 50 and 100 ft. and two heights of eye-level, viz. with *S* standing in a normal position and on a box 32 in. in height.[2]

The same procedure was followed every day. *S* was instructed not to look at the targets while *E* set the variable in a near or far position. At the signal 'Ready,' *S* watched the targets while *E* moved the variable toward the standard at an imperceptible rate. *S* signaled 'Stop' when he judged the targets to be equal in depth. *E* then recorded the measurement from the tape-chain control and set the target for the next trial. *S* made single judgments in a rotating sequence until he had made 10 discriminations under a particular condition. *S* moved off to the side and sat down with his back toward the targets after each trial. Following the complete administration of an experimental condition *S* had a 30-min. rest after which another condition was run. Two conditions were given each morning and each

[2] The average normal height of the eye-level of the 4 *S*s was 14.9 in. above the targets. This resulted in an angle of incidence between *S*'s line of sight and the base-surround of the targets of 1.4° and 0.7° at 50 and 100 ft., respectively. Standing on the box increased the height of *S*'s eye-level another 32 in. Hence the respective angles then became 4.4° and 2.2°. Thus for a given height of eye-level the angle of incidence is decreased with increases in viewing distance.

afternoon. Thus, every condition was repeated twice with an equal number of near and far settings of the variable target presented in each replication.

Results. To determine whether there was a significant constant error (*CE*) in the experiment, the *CE*s were analyzed separately for the near and far settings with respect to their directional signs and their magnitudes. The statistical sign-test was used for this purpose.[3] The hypothesis that there was no difference in the directional signs of these *CE*s was rejected at better than the 0.05-level of confidence. The hypothesis that there was no difference in the magnitudes of these *CE*s could not be rejected at the 0.05-level. On the basis of these findings it was concluded

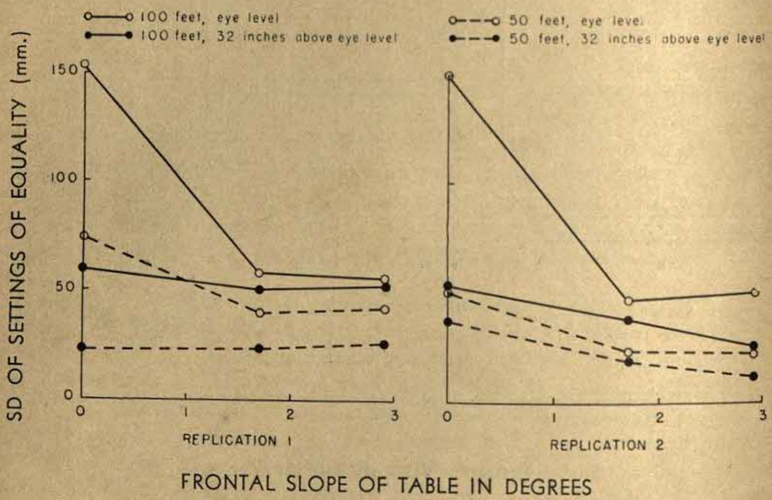


FIG. 2. THRESHOLD OF DEPTH-DISCRIMINATION AS A FUNCTION OF DISTANCE, SLOPE OF TABLE, AND HEIGHT OF EYE-LEVEL

that measures taken over both near and far settings would effectively eliminate all *CE*s, *i.e.* *CE*s obtained by this procedure would be random deviations from zero.

The root mean squares of the equality settings corrected for the obtained (random) *CE*s were calculated as the best estimates of the precision of the settings about physical zero (perfect matching). One such measure was calculated for the data obtained from all *S*s under each replication of a condition. The results are shown graphically in Fig. 2.

Fig. 2 shows the relationship between the standard deviation (*SD*) of equality settings (in mm.) and the frontal-slope of the table for the two distances, two eye-levels, and two replications used. In general an increase in slope is associated with a decrease in *SD*. From the curves presented it is also apparent that increasing the height of *S*'s eye-level decreased the *SD* while increasing viewing distance in-

[3] W. J. Dixon and J. F. Massey, *Introduction to Statistical Analysis*, 1951, 247-253.

creased the *SD*. A comparison of corresponding points on the respective curves for the two replications shows that the *SD* tended to decrease with practice.

Since the measures shown in Fig. 2 are *SD*s and since only one *SD* was obtained under each replication of a condition, it could not be assumed that the measures constitute samples from a normally distributed population with equal means and equal variances. To determine the statistical significance of the differences shown in Fig. 2, all *SD*s shown were transformed to normalized ranks with the aid of tables provided by Fisher and Yates.[4] An analysis of variance was then performed upon the ranked data. A summary of this analysis is shown in Table I.

The two- and three-factor interactions in Table I were evaluated using the four-factor interaction mean square as the error term. Since none of the interactions was statistically significant their mean squares were all assumed to be estimates of the same random errors. A pooled error-term based on all interactions sums of squares was calculated and is shown in Table I. Using this error-term, the *F*-tests for each main effect of distance, height of eye-level, slope of base-surround, and replication were all significant at better than the 0.001-level of confidence. Thus it was concluded that the means of the normalized ranks of the *SD*s for the levels of each main effect differed significantly and that these main effects were statistically independent of one another.

TABLE I

ANALYSIS OF VARIANCE OF NORMALIZED RANKS OF THE *SD* FOR EACH REPLICATION OF EACH CONDITION

Source	*df*	*SS*	*MS*
Distance (*D*)	1	7.6840	7.6840*
Eye-level (*L*)	1	3.8400	3.8400*
Slope (*Sl*)	2	5.2575	2.6288*
Replication (*R*)	1	2.4962	2.4962*
D×*L*	1	.1204	.1204
D×*Sl*	2	.0398	.0199
D×*R*	1	.0054	.0054
L×*Sl*	2	.5028	.2514
L×*R*	1	.0433	.0433
Sl×*R*	2	.6120	.3060
D×*L*×*Sl*	2	.0891	.0445
D×*L*×*R*	1	.0817	.0817
D×*Sl*×*R*	2	.2086	.1043
L×*Sl*×*R*	2	.3722	.1861
D×*L*×*Sl*×*R*	2	.2510	.1255
Total	23	21.6040	
Pooled average	18	2.3263	.1292

* Significant at the 0.001-level of confidence.

In Fig. 3 the *SD*s of the settings of equality are plotted as a function of the angle of incidence between *S*'s line of sight and the base-surround of the targets.[5] These curves clearly show that slight changes in the angle when it is small greatly

[4] R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research*, 1948, 66-67.

[5] The total angle of incidence was a function of as many as four factors: (1) table-slope; (2) average normal height of *S*'s eye-level; (3) height of *S*'s eye-level when standing on the box; and (4) the height of the target's base-surround.

affect the *SD*. They also suggest that the rate of change is greater at a distance of 100 than at 50 ft. In general the *SD* decreases with an increase in the total angle of incidence. Results similar to those in Fig. 2 were found with respect to the effects of increased distance and increased practice.

Discussion. The results of this experiment clearly show that the angular relationships between *S* and the target's base-surround are important factors in the discrimination of depth. The effects are similar whether the angle is caused by the slope of the base-surround or by changes in the height of *S*'s eye-level relative to the base-surround of the target. An increase of either will increase the precision of judgments of the depth. These results support the conclusions of Teichner, Kobrick, and Wehrkamp.[6] They also emphasize the need for a careful control of

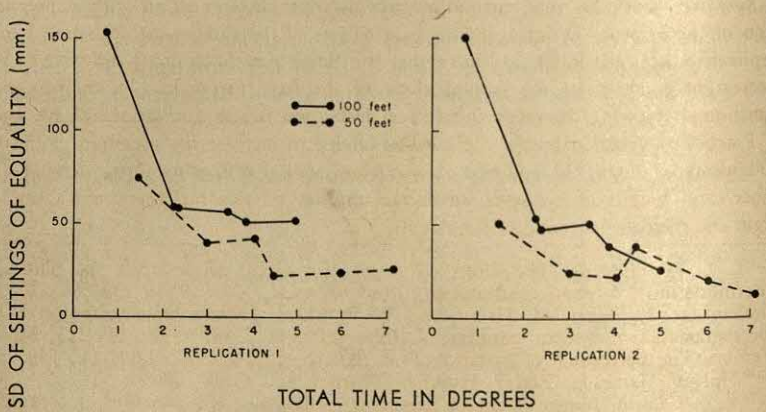


FIG. 3. THRESHOLD OF DEPTH-DISCRIMINATION AS A FUNCTION OF TOTAL ANGLE OF INCIDENCE OF LINE OF SIGHT AND THE BASE-SURROUND FOR TWO DISTANCES

the slope of terrain in discriminating distances out-of-doors. A plausible explanation of the effect of angular relationships between *S* and the base-surround is that it aids in judgments of relative depth by increasing vernier acuity thereby reducing the linear threshold of depth-discrimination. Such an explanation is considered in detail in the study previously mentioned.[7]

The increase in *SD* with an increase in distance found in this experiment is supported by earlier studies.[8] In the present study the visual angle subtended by the target at the nearer distance was twice that at the farther distance; however, most results obtained by previous investigators support the conclusion that other effects

[6] Teichner, Kobrick, and Wehrkamp, *op. cit.,* 193-208.
[7] *Loc. cit.*
[8] M. J. Hirsch and F. W. Weymouth, Distance-discrimination: V. Effect of motion and distance of targets on monocular and binocular distance discrimination, *J. Aviat. Med.,* 18, 1947, 594-600; Teichner, Kobrick, and Wehrkamp, *op. cit.,* 193-208. J. G. Beebe-Center, Leonard Carmichael, and L. C. Mead, Daylight training of pilots for night flying, *Aeronaut. Eng. Rev.,* 3, 1944, 1-10; A. H. Holway, D. A. Jameson, M. J. Zigler, L. M. Hurvich, A. B. Warren, and E. B. Cook, Factors influencing the magnitude of range-errors in free-space and telescopic vision, *Div. of Res., Grad. Sch. of Bus. Adm., Harvard Univ.,* 1945, 1-7.

associated with distance are preponderant.[9] Contradictory results have also been obtained.[10] In the present experiment the angular relationships caused by the height of eye-level contribute considerably to the effects of distance.

The effects of practice that were obtained are all the more significant since preliminary training was given to minimize them. Gibson, in a survey of studies relating perceptual judgments to learning, reports two experiments with stereoptic rangefinders in which the discrimination of stereoscopic depth improved with practice.[11] The results of the present study indicate that perceptual learning occurs in unaided judgments of relative depth-discrimination as well. The possible presence of perceptual learning should not be overlooked even when using experienced *Ss*.

Summary. Four *Ss* were used in a study of depth-discrimination with a modification of the Howard-Dolman apparatus. Analysis of the three variables studied in a replicated factorial design indicates that the linear threshold of equality is: (1) a decreasing function of the frontal slope of the base-surround; (2) an increasing function of viewing distance; and (3) a decreasing function of height of eye-level.

Practice as revealed by the replications tended to increase the precision of the *S*'s judgments of depth. In addition it was concluded that frontal-slope, viewing distance, and height of eye-level affect the angular relationship between *S*'s line of sight and the base-surround of the targets.

---

[9] L. S. Woodburne, The effect of a constant visual angle upon the binocular discrimination of depth differences, this JOURNAL, 46, 1934, 273-286; C. H. Graham, K. E. Baker, M. Hecht, and V. V. Lloyd, Factors influencing thresholds for monocular movement parallax, *J. Exper. Psychol.*, 38, 1948, 205-223; M. D. Vernon, The perception of distance, *Brit. J. Psychol.*, 28, 1937, 1-11, 115-149.

[10] Holway, Jameson, Zigler, Hurvich, Warren, and Cook, *op. cit.*, 1-7.

[11] E. J. Gibson, Improvement in perceptual judgments as a function of controlled practice or training, *Psychol. Bull.*, 50, 1953, 401-431.

Perceptual and Oculomotor Efficiency in Reading Materials in Vertical and Horizontal Arrangements

By Miles A. Tinker, University of Minnesota

The traditional arrangement of written and printed words in our western civilization has been in straight horizontal lines. This became the established alignment centuries ago because the scribes found it convenient to write their characters that way. On the other hand, the Chinese and Japanese have, over the centuries, preferred a vertical arrangement of their written and printed symbols. During recent years the latter have arranged some of their printing in horizontal lines.

Huey was the first to raise the question of whether a vertical arrangement of printed words in columns might produce a more efficient typographical arrangement than the traditional horizontal alignment.[1] From the theoretical point of view he considered that a vertical arrangement should prove more efficient due to elimination of practically all horizontal eye-movements during reading. Use of the vertical as well as the horizontal span of vision should increase the possibility of having more words in clear vision during any fixation-pause.

In this same report, some preliminary experiments are described by Huey. He found that nonsense material could be read aloud as fast in the vertical as in the horizontal arrangement. For sense material the vertical was read from 7 to 10% slower. When read silently the vertical material was read considerably slower. No *S* had had practice with materials in a vertical arrangement. Furthermore he used only 3 to 5 *S*s in different parts of the experiment. Considering the enormous practice with horizontal printing and lack of practice with the vertical, Huey considered that with practice the vertical might eventually be read faster.

For over 50 yr. there has been no additional investigation of reading English materials in a vertical arrangement. In addition to the possible advantages noted above for a vertical arrangement, there are others. Ruediger, in mapping out the field of distinct vision, found that the field varied in different individuals from a 'square-oval,' about twice as long horizontally as wide vertically, to a shorter oblong.[2] Furthermore, he noted that clear vision extended just as far to the left of the fixation as to the right. Thus with an arrangement which permits use of both the vertical and the horizontal visual span, one should be able to perceive more words per fixation than in the ordinary long horizontal lines of print.

In the past no attempt has been made to compare reading of vertical with horizontal materials when *S*s are given systematic practice in reading the vertical. Obviously habit due to practice is an important determining factor in the relative efficiency with which material in the two typographical arrangements can be read.

* Accepted for publication October 19, 1953. This study was made possible by a grant-in-aid from the Graduate School of the University of Minnesota.

[1] E. B. Huey, Preliminary experiments in the physiology and psychology of readings, this Journal, 9, 1898, 575-586.

[2] W. C. Ruediger, Field of distinct vision, *Arch. Psychol.*, 1, 1907 (No. 5), 1-69.

This experiment was planned to investigate the effect of a limited period of systematic, controlled practice in reading vertical materials upon speed of perception and patterns of eye-movement while reading strictly comparable materials in the two arrangements.

Materials and procedure. The first task was to obtain selections of comparable reading difficulty for the initial and final tests of reading speed. Eight passages of 300 words each were selected from an adaptation (fifth grade difficulty) of *Lorna Doone*. These selections were timed while read by 123 college sophomores. Comprehension was checked. The mean reading times for the four passages chosen for the experiment were respectively 43.75, 44.30, 43.40, and 43.10 sec. The standard deviations were similarly comparable. In a like manner, four comparable passages of 72 words each were selected for the initial and final recording of eye-movements during reading.

The selections for the test of speed were typed in pica on 8½ by 11-in. sheets. A 40-space (4 in.) line was employed for the horizontal material with single spacing between lines and double spacing between paragraphs. The vertical material was arranged in four columns per page, each column 4 in. in length, *i.e.* 24 lines with one word per line. The words in each column were aligned at the left. Each of the four comparable selections was set up in both horizontal and vertical arrangements. The four 72-word selections were arranged similarly for the recording of eye-movements. Two of the speeds and two of the eye-movements selected were employed in the initial tests, and the remaining two in the final tests. Both test-materials and *Ss* were rotated systematically to equate the effects of practice.

In addition, twelve 300-word selections were organized in a vertical arrangement for the controlled practice which was to follow the initial tests. Although taken from the same source (*Lorna Doone*), these practice exercises were not equated for difficulty.

Subjects. Ten college women served as *Ss*. For the initial speed-test, a stop watch was used to record time in seconds. All experimenting was done individually under uniform conditions. The eye-movements were photographed with the Minnesota Camera devised by Tinker.[3] Every *S* read a practice exercise before the camera prior to reading each kind of experimental material in order to accustom her to the experimental situation.

After the initial test was completed, successive periods of controlled practice were started. These periods of practice continued for approximately 6 wk. All practice was on reading 300-word selections in the vertical arrangement. Every *S* read a total of 12 selections. The first six were read four times each, the last six, three times each—a total of 42 readings.

Instructions. S was instructed to strive to increase her speed of reading. Directions on how to do this were given, *i.e.* to move the eyes downward along the columns of print, to avoid horizontal movements, to try to perceive three to four words per fixation, and to use as few fixations per column as possible.

At the end of the 42 practice readings, *S* was again tested for speed of reading and eye-movements were photographed while reading horizontal and vertical materials. A questionary checked reactions to reading the vertical print.

[3] M. A. Tinker, Apparatus for recording eye-movements, *Amer. J. Psychol.*, 43, 1931, 115-118.

Results. In interpreting the results it is necessary to keep in mind the relatively short time devoted to practice on the vertical print in comparison to the years devoted to reading materials in horizontal alignment. Results of the initial and final speeds of reading are given in Table I. As expected, the initial testing revealed that the vertical material was read much slower than the horizontal. The vertical was on the average read 24.7 sec. slower than the horizontal, which represents a 50-% retardation. After the practice in reading vertical print, the difference was only 10.9 sec. or 21.8% in favor of the horizontal. This difference, however, is still significant beyond the 1-% level.[4]

The third section of the table reveals that there was no change from the initial to the final testing in reading horizontal print, but the last section shows that there

TABLE I

INITIAL AND FINAL SPEED OF READING FOR HORIZONTAL AND VERTICAL ARRANGEMENTS
(The mean scores are seconds taken to read selections of 300 words by 10 *Ss*.)

Comparison	Mean	*SD*	Diff.	% diff.	*t*
Horizontal: Initial	49.4	11.2	+24.7	+50.0	7.38*
Vertical: Initial	74.1	12.6			
Horizontal: Final	50.0	13.7	+10.9	+21.8	6.02*
Vertical: Final	60.9	12.4			
Horizontal: Initial	49.4	11.2	+ 0.6	+ 1.2	0.23
Horizontal: Final	50.0	13.7			
Vertical: Initial	74.1	12.6	−13.2	−17.8	5.92*
Vertical: Final	60.9	12.4			

* Significant beyond the 1-% level.

was a significant gain of 13.2 sec. or 17.8% from initial to final reading of the vertical print. Examination of the individual data reveals that one *S* read the vertical as fast as the horizontal after the practice although she had a 17-sec. difference on the initial testing. Two *Ss* were only 6 and 7 sec. slower on the vertical after the practice. The group as a whole read nearly one word more per second (4.05 to 4.92) after the practice.

The eye-movements for initial and final testing are given in Table II. Since the reading sample employed for photographing eye-movements was brief (72 words), and since the objective was merely to note trends in oculomotor patterns, no tests of significance were made. An analysis of these data will reveal the differences in oculomotor patterns employed in reading vertical in comparison with horizontal print. On the initial testing there were 13.7 fewer fixations in reading the vertical material or a 22.4% difference. The difference was in the same direction but not as large on the final testing; 7.5 fewer fixations or 14.6%.

The number of words read per fixation reflects the findings just noted. On the initial testing, 28.8% more words per fixation were read in the vertical; and on the final testing, 18.5% more.

Regression-frequency was much less for the vertical print. Furthermore, there was greater improvement in cutting down regressions in vertical reading after the

[4] The *t*-formula (R. A. Fisher, *Statistical Methods for Research Workers,* 1936) for small samples was used to compute significance.

practice. At the initial testing there was 45.9% fewer regressions on the vertical; at the final testing, 59.6% fewer.

The duration of pauses was 42.9% longer for the vertical on the initial testing, and 64.7% longer on the final testing.

The general pattern of the eye-movements reveals that the vertical reading required fewer fixations, fewer regressions, a longer pause duration, and more words were read per fixation. Practice in reading vertical print brought an improvement in number of fixations and regressions in comparison with horizontal reading.

To sum up the trends in Tables I and II, we find: (1) Prior to practice in reading print arranged in columns, the vertical print was read much slower than in the

TABLE II

INITIAL AND FINAL MEASURES OF EYE-MOVEMENTS FOR READING MATERIALS IN HORIZONTAL AND VERTICAL ARRANGEMENTS
(Pause-duration in sec.)

Measure	Mean	*SD*	Diff.	% diff.
No. fixations: Horizontal, initial	61.1	10.1	−13.7	22.4
No. fixations: Vertical, initial	47.4	10.5		
No. fixations: Horizontal, final	51.4	7.3	− 7.5	14.6
No. fixations: Vertical, final	43.9	6.9		
Words per fixation: Horizontal, initial	1.18	.23	+ .34	28.8
Words per fixation: Vertical, initial	1.52	.31		
Words per fixation: Horizontal, final	1.40	.20	+ .26	18.5
Words per fixation: Vertical, final	1.66	.26		
No. regressions: Horizontal, initial	7.4	5.4	− 3.4	45.9
No. regressions: Vertical, initial	4.0	1.8		
No. regressions: Horizontal, final	4.7	2.9	− 2.8	59.6
No. regressions: Vertical, final	1.9	1.9		
Pause-duration: Horizontal, initial	.21	.03	+ .09	42.9
Pause-duration: Vertical, initial	.30	.05		
Pause-duration: Horizontal, final	.17	.04	+ .11	64.7
Pause-duration: Vertical, final	.28	.04		

customary horizontal arrangement. (2) After a relatively small amount of practice there was a significant gain in reading the vertical material although it was still read significantly slower than print in horizontal alignment. (3) These changes were reflected in the eye-movement records where there was an improvement in fixation- and regression-frequency.

The results of the questionary, which were obtained just before the final tests, revealed the following views of our *Ss* toward reading vertical materials. All of the *Ss* agreed that the practice made the reading easier; five considered that the vertical continued to be more difficult to read; seven seldom had to reread any parts of the vertical to get the thought; eight seldom noted any horizontal movements of the eyes in vertical reading; and five considered that the vertical reading progressed smoothly. The *Ss* did have difficulty, however, in seeing punctuation marks in the vertical arrangement.

The results of this study seem to imply that habit acquired through years of reading materials in horizontal arrangement was the important factor in determining the relative reading rates in the two arrangements. Throughout the experiment the mean rate remained constant for the horizontal print but improved markedly for

the vertical even though the practice on it was small. Horizontal reading skills of college students are, of course, well established. Even at the end of the training, development of skill in vertical reading was only in the preliminary stages. If, however, a reader were originally trained to read materials in a vertical arrangement, the vertical might be as efficient or more efficient than the horizontal. This inference is born out by the results obtained in reading Chinese in vertical and horizontal arrangements. Employing Chinese college students in America, both Shen and Tu found that the vertical materials were read significantly faster.[5] The greater efficiency for the vertical was attributed to the influence of long established reading habits.

Certain factors should be considered in any further investigation of vertical reading. A system of punctuation easily discerned by the reader should be employed. The optimal size of type should be employed so that as large a group of words as possible could be covered in one fixation. Perhaps some attention should be given to grouping words in thought-units. The possibility of placing more than one word on a line when the words are short should also be considered. There should be an investigation to determine the optimum length of columns of words and the best spacing between columns.

To a large degree the reading of print arranged in vertical columns will remain a matter of theoretical interest. Nevertheless, in terms of the alleged advantages for increased efficiency in perceiving larger groups of words per fixation due to the shape of the area of clear vision, and the elimination of most horizontal eye-movements, vertical reading should receive some consideration. Furthermore, this study has demonstrated marked improvement in vertical reading from a small amount of practice.

Important obstacles to use of vertical print are: (1) tradition, (2) the marked variation in length of English words, and (3) lack of knowledge concerning optimal typographical arrangements for vertical printing.

In view of the possible efficiency of vertical reading and the resistance to change from the traditional horizontal arrangement, it might be possible to work out a compromise between the two. In fact, Andrews has proposed a typographical arrangement called 'square span' in which the material is printed in double-line block.[6] For example:

Thinking consists partly in	making use of what	has been learned before

Andrews considers that 'square span' printing should aid reading speed and comprehension by effectively employing both the horizontal and vertical visual span and by grouping the words into thought-units. In his experiment, Andrews discovered a slight advantage for the 'square span' printing. Suspecting that any advantage for the square span lay in the grouping of words into thought-units, North and Jenkins compared square span with normal horizontal printing divided

---

[5] Eugene Shen, An analysis of eye movements in the reading of Chinese, *J. Exper. Psychol.*, 10, 1927, 158-183; H. T. C. Tu, The effect of different arrangements of the Chinese language upon speed and comprehension of silent reading, *J. Genet. Psychol.*, 38, 1930, 321-337.

[6] R. B. Andrews, Reading power unlimited, *Texas Outlook,* 33, 1949, 20-21.

by extra spacing into thought-units called "spaced units."[7] Measures of speed and comprehension were obtained. Their hypothesis was confirmed when they found that the spaced unit printing was superior to the square span and the normal horizontal arrangement. The square span printing was slightly less efficient than the normal horizontal arrangement. The question still remains as to whether with sufficient practice the square span can be read as fast or faster than normal horizontal printing. In any case, the use of either square span or spaced unit printing would probably be useful only in reading situations where the message is brief and where speed of perception is at a premium.

Summary. The purpose of this experiment was to investigate the effect of a limited period of practice in reading materials arranged in vertical columns upon speed and patterns of eye-movements while reading comparable materials in vertical and in normal horizontal arrangements. Equated passages of easy material were employed for initial and final tests. Twelve similar passages were employed for the practice materials. Each of the practice selections was read in vertical arrangement three or four times over a period of 6 wk. which elapsed between the initial and final tests. The following results were obtained.

(1) At the initial testing, the vertical reading was 50% slower than the horizontal, but after practice it was only 21.8% slower. Speed of reading the horizontal did not change from initial to final testing but there was a significant gain of 17.8% in the vertical reading.

(2) The records of the eye-movements on the initial test revealed that the vertical reading required fewer fixations, fewer regressions, and longer pause-durations. Practice yielded an improvement in number of fixations and regressions in vertical reading in comparison with the horizontal.

(3) The results imply that long established reading habits produced the superiority of horizontal over vertical reading.

(4) The marked improvement in efficiency of vertical reading with slight practice suggests that the vertical arrangement could become as efficient or more efficient than the horizontal with extensive practice in reading vertically arranged materials.

(5) With the possible exception of 'square span' printing, which is a compromise between the vertical and horizontal arrangements, the efficiency of vertically arranged printing is apt to remain largely a matter of theoretical interest.

[7] A. J. North and L. B. Jenkins, Reading speed and comprehension as a function of typography, *J. Appl. Psychol.*, 35, 1951, 225-228.

A MODIFICATION OF THE METHOD OF EQUAL-APPEARING INTERVALS

By J. P. GUILFORD and HARVEY F. DINGMAN,
University of Southern California

One of the obvious defects of the method of equal-appearing intervals has been the 'end effect,' which results in seriously truncated frequency-distributions for stimuli near the extremes of the list. This, in turn, results in difficulties in the estimations of central tendency and variability. It may have other, less obvious, effects in the form of distortions of the scale so that we do not have, after all, an equal-interval-scale.

We propose here a procedure that is aimed at elimination of the 'end effect.' It was applied to a series of lifted-weights, with treatment of the data by the method of successive categories to determine how well equality of intervals was achieved. An incidental result is a test of the applicability of Fechner's law to lifted-weights.

Procedure. The principle of the device designed to overcome the end effect was to provide the observer (*O*) with categories beyond the range of the nine that were provided to cover the range of stimuli being judged. At either end of the scale, *O* could let his judgment run over into one of the extra categories whenever he felt a weight to be more extreme than it should for placement in the end category. There were three extra categories at either end, making 15 usable classes into which to place stimuli.

A series of 21 weights, loaded ointment boxes, was selected, which roughly varied geometrically. The lightest weight was 50 gm. and the heaviest 100 gm. A small range is desirable, with small increments, otherwise the dispersion of any one stimulus over the categories is likely to be very narrow. Anchor stimuli of 50 and 100 gm. were placed in definition of categories numbered 4 and 12, respectively. It was thought that without such anchors the weights would tend to be distributed more widely over all 15 categories and the end effect would after all be in evidence. The anchors were lifted only when *O* wished to refresh his impression of what weights in Categories 4 and 12 should feel like. We recognized that the anchors might have some distorting effects on the scale, but this could be detected by the process of scaling the categories.

The list of weights is described in terms of grams in Table III. In the experimental series, each weight appeared three times, so that *O* placed 63 weights in one set of judgments. Fifteen *O*s provided 2 sets of judgments each and 5 others one set each, making 105 placements for each stimulus-weight. The *O*s were mostly graduate students in psychology.

Results: (1) Frequency-distributions. In Table I we have presented frequency-distributions for eight selected stimuli; three of which are at the low end of the list, three at intermediate levels, and two at the high end. The sums of all frequencies (for the 21 weights) in the bottom row show to what extent the extra categories were used. A total of only 11 judgments appears in the 3 lowest categories, in-

* Accepted for publication August 19, 1954.

dicating they were used very little. A total of 100 judgments appears in the 3 highest categories. The reason for this asymmetry in the use of extra categories is not obvious. In general, the higher categories were used relatively more than the lower ones, resulting in an average judgment of 8.7 for all weights combined, whereas 8.0 is the midpoint of the 15-step range.

Two distributions, for weights 53.5 and 57.5, show some hint of truncation. Other distributions, including those for weights 51.5, 97.5, and 100.0, each shows what may be called a 'precipice' at the boundary of the category having the anchoring stimulus. Thus, the end effect has not been entirely overcome, but it has been con-

TABLE I.

SAMPLE FREQUENCY-DISTRIBUTIONS OF CATEGORICAL JUDGMENTS FOR EIGHT SELECTED WEIGHTS AND SUMS OF THEIR FREQUENCIES

W	Category														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
50.0			9	27	35	17	12	4	1						
51.5		1	1	30	33	27	7	2	4						
53.5				16	37	15	23	11	2	1					
57.5				10	10	28	29	18	8	1	1				
64.0				1	7	19	26	23	19	8	2				
79.0						2	5	16	13	28	30	7	4		
97.5								1	5	7	15	59	12	6	
100.0									4	5	11	56	22	5	2
Σ_f		1	10	100	187	217	241	262	278	266	284	259	78	16	6

siderably reduced. If Categories 4 and 12 had caught all the judgments that went beyond them, while retaining the ones they already have, there would have been several serious truncations.

(2) Scaling the categories. The scaling of the categories was carried out according to procedures described by Guilford.[1] A single value is thus obtained for each category within which judgments have been placed. The resulting values are given in the second column of Table II. Category 1 could not be scaled because no stimuli were assigned to it. Categories 2 and 15 are scaled with some doubt, because a limited number of stimuli reached that far.

Of most interest to us at this point is the question of whether the intervals of the categories are equal. The interval-widths are not to be found directly by taking successive differences beween the values of neighboring categories. Assuming that the value of each category is at the midpoint of its interval, we arrive at the interval-widths given in Table II. These are plotted as a function of category scale-values in Fig. 1, to show the systematic variation in category-widths. The maximal widths come at Categories 4 and 12. This could have been predicted from knowledge of the unduly large frequencies in those categories. The inequality of intervals pertains not only to Categories 4 and 12 but also to others all along the line. How much of this is due to the use of anchoring stimuli we are unable to say. They must have had much to do with it.

(3) Scaling. The main objective of the method of equal-appearing intervals is to

[1] J. P. Guilford, *Psychometric Methods*, 2d ed., 1954, 237-241.

arrive at psychological scale-values of the weights. We derived scale-values for the 21 weights as in the usual treatment of equal-appearing interval-judgments, assuming equality of intervals, and also scale-values based upon the obtained category values

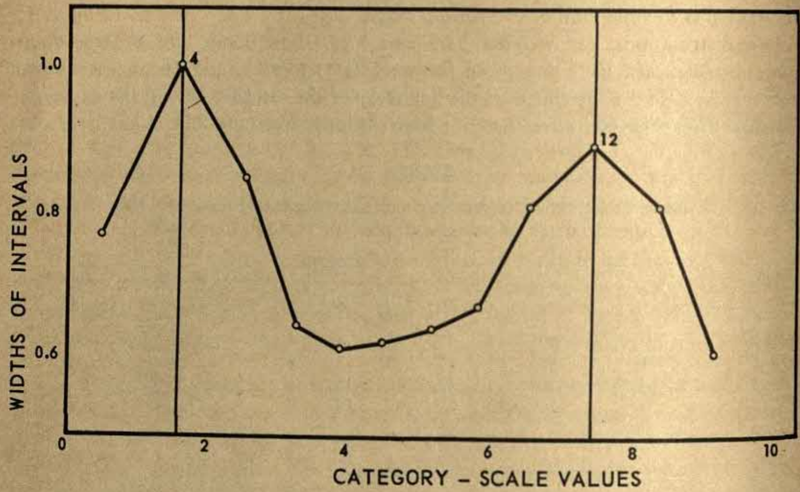


FIG. 1. GRAPHIC ILLUSTRATION OF THE TREND IN CATEGORY WIDTH IN RELATION TO CATEGORY SCALE-VALUES

Maximal widths came at Categories 4 and 12, which had the anchor stimuli.

TABLE II

QUANTITATIVE INFORMATION REGARDING THE INTERVALS CORRESPONDING TO THE CATEGORIES

Interval	C-scale value	Interval width	Adjusted category values
1	—	—	—
2	.00	—	1.6
3	.46	.77	2.2
4	1.55	1.02	3.8
5	2.50	.85	5.1
6	3.25	.65	6.2
7	3.89	.62	7.1
8	4.50	.63	8.0
9	5.15	.65	8.9
10	5.80	.68	9.8
11	6.51	.82	10.8
12	7.43	.90	12.1
13	8.30	.82	13.4
14	9.07	.62	14.5
15	9.54	—	15.1

given in Table II. There is also interest in measures of variability of the different stimuli, which we have also computed in both ways. The resulting means and standard deviations are given in Table III.

The 21 scale-values from the method of equal-appearing intervals, M_e, are not directly comparable with those from the method of successive categories, M_c, owing to differences in unit and origin. The two sets of values actually intercorrelate 0.999. Applying a linear transformation to the M_c values, using the equation $M_{ce} = 1.423 M_c + 1.56$, we arrive at the values in Column 4 of Table III. The M_{ce} values differ from the M_e usually only in the second decimal place. Thus, it can be seen that considerable distortion of the equal-appearing interval-scale can occur without

TABLE III

COMPARISON OF MEANS (M) AND SD OF THE WEIGHTS (W) DERIVED BY THE METHOD OF EQUAL-APPEARING INTERVALS AND THE METHOD OF SUCCESSIVE CATEGORIES

W	M_c	M_e	M_{ce}	SD_c	SD_e	SD_{ce}
50.0	2.46	5.11	5.06	1.06	1.31	1.49
51.5	2.61	5.27	5.27	.98	1.29	1.38
53.5	3.06	5.87	5.91	.97	1.38	1.37
55.5	3.28	6.14	6.22	.91	1.29	1.28
57.5	3.60	6.65	6.69	1.03	1.47	1.45
59.9	3.72	6.80	6.85	.99	1.48	1.40
61.5	4.06	7.31	7.34	.96	1.47	1.35
64.0	4.21	7.51	7.56	.96	1.49	1.35
66.0	4.61	8.16	8.13	.96	1.50	1.35
68.5	4.74	8.34	8.30	.98	1.50	1.38
71.0	5.04	8.79	8.74	.91	1.41	1.28
73.5	5.26	9.12	9.04	1.00	1.49	1.41
76.0	5.48	9.48	9.36	.91	1.38	1.28
79.0	5.79	9.88	9.80	1.07	1.54	1.50
81.5	5.89	10.05	9.95	.90	1.28	1.26
84.5	6.16	10.40	10.33	1.03	1.40	1.45
87.5	6.57	10.95	10.90	1.12	1.44	1.58
91.0	6.85	11.30	11.30	.99	1.26	1.39
94.0	7.11	11.63	11.68	.90	1.07	1.26
97.5	7.25	11.77	11.88	.91	1.14	1.28
100.4	7.47	12.05	12.19	.91	1.12	1.28
M	5.01	8.69	8.69	.973	1.37	1.37
SD	1.52	2.17	2.17	—	—	—

affecting the psychological means of stimuli appreciably. How far this result can be generalized we do not know.

The effect of the scale-distortion upon measures of variability, however, is another story. Standard deviations (SD) were computed for the 21 stimuli, using their frequency-distributions on the two scales, with results under SD_e and SD_c in Table III. Even after a linear transformation of SD_c to values with unit comparable with that for SD_e, we find many discrepancies, as seen in the column under SD_{ce}. The correlation between SD_e and SD_c is 0.52, which, although significant beyond the 0.05-level, is far from what it should be. Thus, distortions of the interval-scale in the method of equal-appearing intervals can seriously interfere with good estimates of variability of stimuli. How far this result can be generalized we do not know.

(4) Test of Fechner's Law. The functional relationship between the scale-values for the weights and logarithms of stimulus-values is definitely linear. Within the relatively short range of 50 to 100 gm. the fit is excellent, with a correlation of

0.999. Although this is not a very decisive test, it is of particular interest because of two previous studies, in which the method of fractionation and the method of constant sums were used, in which the Fechner law was not supported.[2] We could not test the laws that were indicated in those studies, since this requires psychological values on a ratio scale, which we do not have from the methods using categorical judgments. Taking the obtained result in this study in conjunction with results in the previous studies, however, we are inclined to the conclusion that the form of psychophysical law is a function of the psychophysical method used. This means that scaling by different methods does not necessarily yield psychological values that are linearly related.

Summary and conclusions. To counteract the end effect in judgments by the method of equal-appearing intervals, three extra categories of judgment were provided beyond the usual terminal categories, which were marked by anchoring stimuli representing the extremes of the list of stimuli being judged. Twenty-one stimulus-weights were judged 105 times each. Scale-values and measures of dispersion were computed by the usual method, assuming equality of intervals. The categories were scaled by the method of successive categories and these scale-values were used as a basis for computing other estimates of means and standard deviations of stimuli.

The end effect, as indicated by truncated distributions, was very materially reduced but not entirely eliminated.

The category widths were systematically in error, with the largest intervals coinciding with the categories having the anchoring stimuli.

The distortion of the interval-scale had no effect upon accuracy of computed psychological means of stimuli but had serious effects upon the accuracy of standard deviations.

Although inclusion of extra terminal categories does not entirely eliminate the end effect, their use is recommended. It is also recommended that scaling of intervals by the method of successive categories be done routinely, where good estimates of variability of stimuli are desired. Although some anchoring seems desirable, the distorting effects of anchors should be studied, and, if possible, their effects should be minimized.

The results incidentally support Fechner's Law, within the range of weights used, in contrast with other laws recently reported when ratio-judgment methods are used.

---

[2] J. P. Guilford and Harvey F. Dingman, A validation study of ratio-judgment methods, this JOURNAL, 67, 1954, 395-410; R. S. Harper and S. S. Stevens, A psychological scale of weight and a formula for its derivation, this JOURNAL, 61, 1948, 343-351.

APPARATUS

AN ELECTRONIC CIRCUIT FOR THE MEASUREMENT OF THE GALVANIC SKIN RESPONSE

By ROBERT C. NICHOLS, V. A. Hospital, Houston, Texas, and THOMAS DAROGE, University of Kentucky

Apparatus for the measurement of the *GSR* may be divided into two main groups, depending on whether alternating current (AC) or direct current (DC) is applied to *S*. AC offers two major advantages over DC. (1) It eliminates polarization, permitting the use of dry silver electrodes; and (2) it makes practical the use of electronic amplifiers which increases the sensitivity beyond that obtained by the usual DC-bridge. Amplifiers also reduce the size of the current that must be applied to *S*, thereby reducing the possibility of disturbing the effect by the act of measuring it.

In preliminary experiments it was found that both the resistance-level of the skin and the relative *GSR* decrease with increasing frequency. The decrease of the resistance-level can be roughly approximated by a negative exponential equation. The *GSR* also decreases with frequency hence there is only a small indication left at 1000 ~.

Although the indication of the *GSR* is less with AC than with DC, the response at 60 ~ is quite satisfactory. The basic resistance of the skin at this frequency was found to be about one-half of that obtained with DC, and it may be assumed that a corresponding decrease in the size of the *GSR* also occurs. In spite of the smaller indications obtained, it was felt that the other advantages of AC justified its use.

Since the *GSR* consists of a relatively small change in a large static level of skin resistance, it is necessary to cancel a large portion of the initial resistance-level to obtain an accurate measure of the change that is of interest. This is done easily in DC-apparatus by means of a bridge. AC-bridges, on the other hand, are both elaborate and expensive, since it is necessary to balance both the resistive and reactive components of the impedance. This balancing is quite time-cosuming, and since *GSR*'s last only a matter of seconds, the usual AC-bridge is impractical.

The AC-circuit devised by Grant offers an ingenious solution to the balancing problem.[1] Grant used AC across *S*, but rectified this potential,

[1] D. A. Grant, A convenient alternating current circuit for measuring *GSR*'s, this JOURNAL, 59, 1946, 149-151.

i.e. changed it into DC, before balancing it against another DC-potential. By this means Grant was able to obtain the advantages of AC across *S* as well as the ease of balancing afforded by DC. The major disadvantage of Grant's apparatus is that it provides no means of amplifying the small potentials involved in the *GSR*, and resorts to the use of a sensitve galvanometer to measure them.

The instrument to be described here incorporates the principle of Grant's apparatus into an electronic circuit. The vacuum-tube circuit allows the use of an inexpensive milliammeter and provides sufficient output for any type of automatic recorder that might be available. It offers the additional advantage of being sensitive and stable enough for research, yet sturdy enough for classroom and student use. It is impossible to damage the meter by any position of the controls or by shorting the electrodes.

Circuit. A circuit-diagram of the apparatus is shown in Fig. 1. For convenience of exposition, the circuit may be considered as being divided into the six stages marked off by dotted lines in the diagram: subject, amplifier, rectifier, meter, power-supply, and shocker.

S's circuit consists of his electrodes mounted in series with a balancing resistance-network controlled by SW-1 and R-13. A small voltage from the filament supply is impressed across this circuit and is controlled by R-2. The resistance-values of the balancing network are such that they can be set to a value of 10 times that of *S*. This causes a voltage division such that 1/11 of the potential from R-2 is developed across *S*. Changes in *S*'s resistance will now produce changes in the potential across him which are approximately proportional to the percentage of his change.

The voltage developed across *S* is amplified by a two-stage resistance-coupled amplifier which composes the amplifier-stage. The amplifier has a gain of about 400.

The output from the amplifier-stage is rectified by K and changed into smooth DC by the filter, C-5, C-6, L-1. A DC-voltage is thus obtained which varies in proportion to the percentage change in *S*'s resistance.

This voltage is applied to one grid of the difference amplifier, V-2, which composes the meter stage. The other grid is supplied with a constant comparison voltage from the voltage-divider, R-26-R-27. The voltage between the two plates of this tube is proportional to the amplified difference of the grid voltages. A voltmeter connected to the plates then gives a reading which is proportional to the percentage change in *S*'s resistance. R-23 and R-24 allow for two meter ranges. The values given are correct for 25% and 50% ranges, respectively. The output terminals provide for the connection of additional meters or an automatic recording device. The output voltage changes approximately 1.5 v. for each percentage change in *S*.

The power supply produces 150 v. which is regulated by a VR-tube. Both the AC-voltage across *S* and the DC comparison-voltage of the bridge circuit are taken from non-regulated parts of the power supply, while the amplifiers are supplied from a regulated source. The instrument is thus insensitive to line-voltage variations, since any change will affect both parts of the bridge equally.

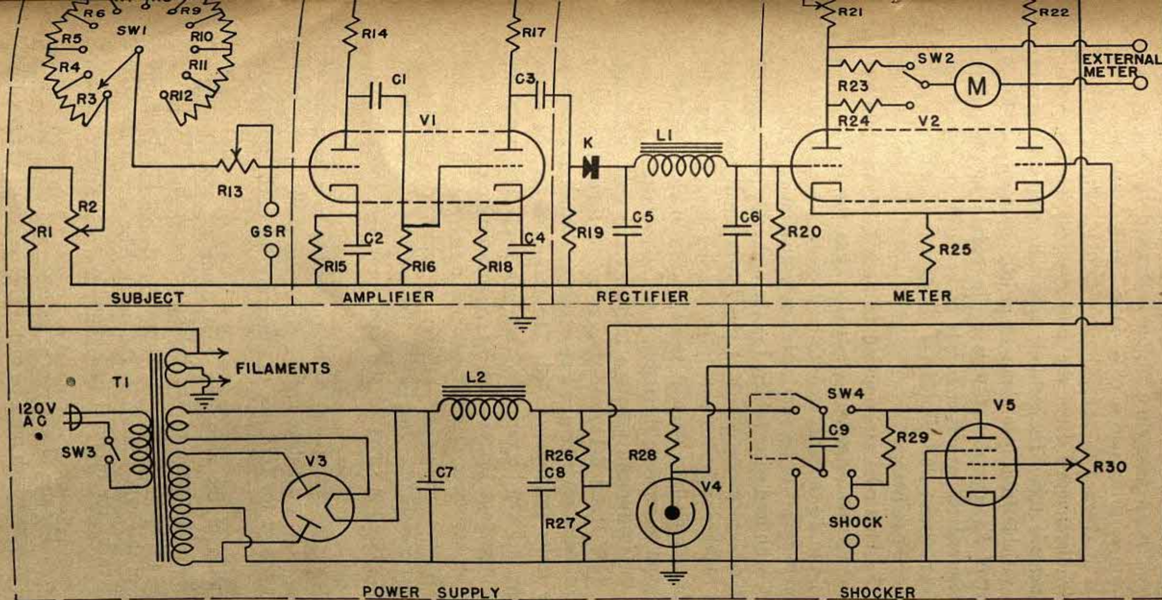


FIG. 1. CIRCUIT DIAGRAM OF THE APPARATUS

Except where otherwise noted, all resistors are 0.5w. and all parts have 10-% tolerance.

C-1 = .25 mfd. 200 v.
C-2 = 20 mfd. 25 v.
C-3 = .25 mfd. 200 v.
C-4 = 20 mfd. 25 v.
C-5 = 2 mfd. 200 v.
C-6 = 4 mfd. 200 v.
C-7 = 8 mfd. 450 v. electrolytic
C-8 = 8 mfd. 450 v. electrolytic
C-9 = 2 mfd. 450 v. electrolytic

(Thordarson T20C50) T20C50)
L-2 = 16 Hy. 50 ma. (Stancor C1003)
K = Germanium Diode
M = 0-50 microammeter
R-1 = 100 Ω
R-2 = 20 Ω linear potentiometer
R-3, R-4, R-5, R-6, R-7,

R-11, R-12, 100,-000 = Ω precision resistors
R-13 = 100,000 Ω linear wire wound potentiometer
R-14 = 47,000 Ω
R-15 = 2,200 Ω
R-16 = 100,000 Ω
R-17 = 47,000 Ω
R-18 = 2,200 Ω
R-19 = 47,000 Ω

potentiometer
R-22 = 220,000 Ω
R-23 = 1,500,000 Ω 1% precision resistor
R-24 = 750,000 Ω 1% precision resistor
R-25 = 6,800 Ω
R-26 = 1,000,000 Ω
R-27 = 47,000 Ω
R-28 = 5,000 Ω 5 w.
R-29 = 6,700 Ω
R-30 = 100,000 Ω linear

T-1 = 120/470 CT/6.3 CT/5 (Stancor PC-8401)
SW-1 = SP 10 T
SW-2 = SPDT
SW-3 = SPST
SW-4 = DPDT spring return switch
V-1 = 6SN7
V-2 = 6SN7
V-3 = 5Y3
V-4 = OD3/VR150

Shocker. The shocker-circuit is not an essential part of the apparatus, but since electric shock is so commonly used as a stimulus for the *GSR*, it was included in the circuit. The shocker-circuit is an adaptation of one suggested by Mount and Lehner.[2] A DPDT switch with a spring return (SW-4) normally connects C-9 to the power supply and keeps it charged to approximately 300 v. When SW-4 is thrown, C-9 is momentarily disconnected from the power supply and is placed across *S* in series with a pentode tube. The pentode tube determines a constant current through *S* which is variable from 0 to about 15 m.amp. by means of R-30. This current is independent of *S*'s resistance. C-9 is discharged in approximately 0.01 sec. by R-29, and is returned to the power supply to be recharged when the switch is released. The shocker-circuit thus provides a constant current for a constant time-interval regardless of *S*'s resistance.

Electrodes. Since one of *S*'s electrodes is grounded in both *GSR* and shocker-circuits, it is possible to use a common ground electrode. Thus the *GSR* can be measured and the stimulus can be administered through only three electrodes. Silver pieces (25¢ coins) were found to be very satisfactory and inexpensive electrodes. Three of these pieces, with lead-wires soldered to the underside, are mounted on top of small sections of rubber sponge, which is glued to a small board. *S* is connected to the *GSR* and shocker-apparatus simply by placing his palm on the silver pieces where it is held securely by a strap. A small amount of electrode paste placed on each piece was found to improve the stability of the contact.

Calibration. To calibrate the instrument, a known resistance is substituted for *S* and the balancing network (SW-1 and R-13) is adjusted to 10 times this value. R-2 is then adjusted until the meter reads zero. *S*'s resistance is then decreased by a known percentage, and the meter reading is noted. The magnitude of the meter deflection for a given change in *S* can be varied by changing R-21. The meter can thus be calibrated to read percentage decrease of *S*'s resistance. This calibration will hold for all values of his resistance. Theoretically the meter should not read linearly with change in resistance because of the way in which the voltage is divided between *S* and the series-resistance. In practice, however, it is almost linear, partially because of the compensating effect of the non-linear vacuum-tube characteristic.

Calibration in percentage change was found to be preferable to a measure of absolute change of resistance. The percentage change eliminates some of the wide individual differences in the size of the *GSR* due to differences in initial resistance-level.

Operation. For the operation of the apparatus, *S* is connected and the meter balanced somewhat above zero by means of the balancing network (SW-1 and R-13). The actual resistance of *S* at any time is the value set on the balancing network less the percentage of this value shown by the meter. The size of the *GSR* in percentage decrease is given by the maximal post-stimulus meter reading minus the pre-stimulus reading.

Transformation of the scores. The use of analysis of variance and other statistical designs involves certain assumptions which must be met if the results of the analysis

[2] G. E. Mount and G. F. J. Lehner, The use of certain electronic tubes in the psychological laboratory, this JOURNAL, 61, 1948, 247-258.

are to be valid. As Haggard has stated,[3] ideally data on the *GSR* should possess or approximate the following characteristics: (a) additivity, (b) normality, (c) homogeneity of the variances, (d) independence of the means and variances, and (e) randomness. It has long been recognized that the usual measure of the *GSR*, in terms of ohms-change, is inadequate when evaluated by these criteria, and numerous transformations have been tried in an attempt to find an appropriate measure. The transformation of *GSR*-scores into log units of conductance change has been found by Haggard to best approximate these criteria.

The present writers have found the most satisfactory unit to be the logarithm to the base 10 of the conductance change in micromhos plus 1 $\mu\mho$ times 100. Although this is not exactly the log conductance change, it will hereafter be called that for simplicity of reference.

A change of 1 $\mu\mho$ in conductance was made before the logarithm was taken to make zero scores more meaningful. A zero score is one in which the stimulus is followed by no increase in conductance. If these scores are counted as zero changes in conductance, the logarithm will be negative infinity. Since 1 $\mu\mho$ has been added, the logarithm of these scores is, however, itself zero. The use of log $X + 1$ is suggested by Edwards when the logarithmic transformation of data containing zero counts is made.[4]

The raw data obtained from the instrument are composed of three measures for each response: (1) the basic level of resistance in ohms, (2) the pre-stimulus resistance in percentage of the basic level, and (3) the minimal post-stimulus resistance in percentage of the basic level. The transformation of these three measures into a single log score of conductance change involves the solving of the following formula:

$$G = \log \left[\frac{100}{R(100 - B)} - \frac{100}{R(100 - A)} + 10^{-6} \right] .10^{6}$$

Where G = log conductance change; R = basic resistance level in ohms; A = pre-stimulus resistance in percentage of R; and B = post-stimulus resistance in percentage of R. Since this formula is quite clumsy to use, a short-cut method of computation is needed when a large number of scores is to be transformed. The writers have found that the scores can be transformed with little difficulty by means of the nomogram shown in Fig. 2.

Use of the nomogram. In this nomogram, Scale A represents the initial reading of the meter and Scale B represents the maximal post-stimulus deflection of the meter. The basic resistance-level of *S* as shown by SW-1 and R-13 is represented by Scale R. The log conductance change is represented by Scale G.

To use the nomogram, the points on Scales A and B corresponding respectively to the pre- and post-stimulus meter-readings are connected by a ruler and the reading is noted at the intersection of the ruler with Scale K. This value is then transferred to Scale L and the ruler is placed between this point and the value corresponding to the basic level of resistance on Scale R. The log score of the conductance

[3] E. A. Haggard, On the application of analysis of variance to GSR data: I. The selection of an appropriate measure, *J. Exper. Psychol.*, 39, 1949, 378-392.

[4] A. L. Edwards, *Experimental Design in Psychological Research*, 1950, 203.

change can then be read from Scale G where it is crossed by the ruler.

In the example shown by the broken lines in Fig. 2, the basic resistance-level shown by SW-1 and R-13 is 22,000 Ω. The pre-stimulus meter-reading is 10% and the maximal post-stimulus deflection is to 17%. A straight line is drawn from point 10 on Scale A through point 17 on Scale B until it intersects Scale K. The reading of 95 is transferred from Scale K to Scale L and this point is connected

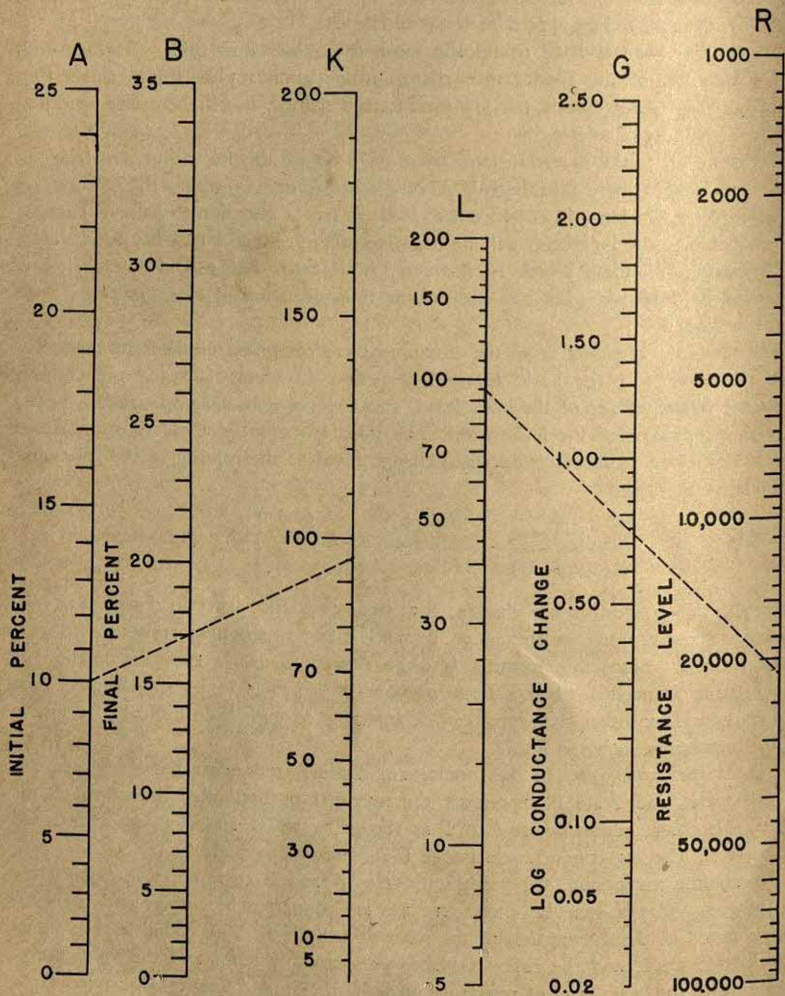


FIG. 2. NOMOGRAM FOR THE TRANSFORMATION OF THE RAW *GSR* SCORES INTO LOG UNITS OF CONDUCTANCE-CHANGE

Example: A = 10%; B = 17%; R = 22,000 Ω; and G = 0.73. Connect A and B to locate K. Transfer the reading on K to L and then connect L and R to locate G.

by a straight line with point 22,000 on Scale R. This line intersects Scale G at 0.73, which is the log score of the conductance change. In practice, these scores are multiplied by 100 to remove decimals.

Summary

An apparatus for measuring the *GSR* has been described which involves passing a small AC through *S*, amplifying the voltage with electronic amplifiers and changing it into DC for comparison with a standard voltage in a bridge circuit. The apparatus includes a circuit for administering a shock-stimulus of a constant current for a constant time-interval. Also described is a nomogram for simplifying the transformation of *GSR* scores into log units of conductance change.

THE NEW YORK UNIVERSITY TACHISTOSCOPE

By LAWRENCE KARLIN, New York University

The Dodge mirror-tachistoscope[1] has been modified in recent years by the use of gas tubes which are switched either electronically[2] or by mechanical relays[3] to present stimuli against a background of constant illumination.

The present article describes a tachistoscope in use at New York University. Fig. 1 shows a plan view of the tachistoscope box with the lid

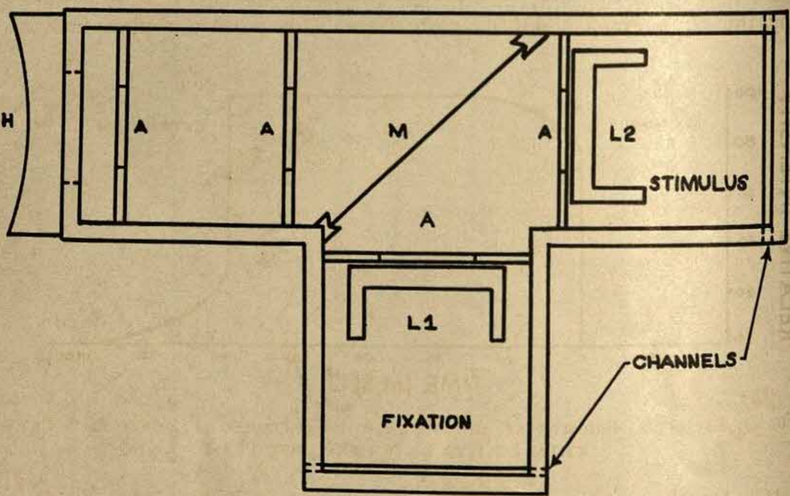


FIG. 1. PLAN VIEW OF MODIFIED DODGE MIRROR-TACHISTOSCOPE
A, aperture; L, lamps; H, viewing hood; M, mirror.

removed. The box is 8 in. in height and made of ⅜-in. plywood. Two horizontal channels are located at the end of each alley. Stimulus-materials are inserted in the channels in the stimulus-compartment. The other compartment furnishes the fixation-stimulus. The T-shape of the present tachistoscope makes it more convenient to adjust the horizontal position of the stimulus-materials than was true of the original square-shaped Dodge tachistoscope box. To facilitate uniform illumination of the stimulus-

[1] Raymond Dodge, An improved exposure apparatus, *Psychol. Bull.*, 4, 1907, 10-13.
[2] J. E. Kupperian and Edwin Golin, An electronic tachistoscope, this JOURNAL, 64, 1951, 274-276.
[3] J. G. Merryman and H. E. Allen, An improved electronic tachistoscope, this JOURNAL, 66, 1953, 110-114.

material, the interiors of these compartments are painted white. The rest of the box interior is painted black. Adjustable stops, located at A, eliminate distracting contours. The edges of the stops bounding the stimulus- and fixation-compartments are lined up optically to eliminate movement of the contours, and they provide an optically more comfortable boundary than the stop in the common channel.

The lamps located as shown in Fig. 1 are 4-in., 5/8 in. dia., cold-cathode, mercury-argon gas-discharge lamps with a striking potential of about 600 v. and a running potential of about 200 v. This type of lamp has been used in the present application both with and without a phosphor coating. Without the phosphor coating the lamp emits a blue light. With the coating the lamp emits a 4500° white light. This particular phosphor was

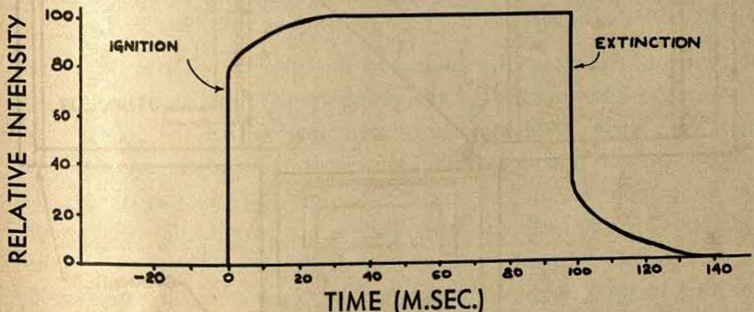


FIG. 2. RELATIVE BRIGHTNESS OF THE PHOSPHOR-COATED FLUORESCENT LAMP WHEN EXCITED BY A 100-M. SEC. PULSE

selected by the writer after trial with various phosphors because of its rapid decay characteristics. These characteristics were determined by flashing the lamp and exciting a phototube, the output of which was measured by a Type 304A Dumont oscilloscope. Fig. 2 shows a copy of the curve so obtained. The curve rises instantaneously to about 78% of peak intensity, after 11 m.sec. to about 90% and by 22 m.sec. has virtually completed its rise. It decays instantaneously to about 30% of peak intensity, to 11% in 11 m.sec. to 5% in 22 m.sec. and is virtaully zero by 33 m.sec.

While the phosphor after-glow has been considered prohibitive in the present application, the writer's experience suggests that phosphors can be selected with decay rates so fast that they will not emit an after-glow of duration long enough and intensity sufficient to interfere with most tachistoscopic work.[4] Another fact to consider is that any stimulus-contours that might ordinarily be visible against a

[4] Merryman and Allen, *op. cit.*, 113.

dark field as a result of such after-glow are rapidly washed out by the ignition of the fixation-lamp that superimposes a uniformly bright field over the stimulus-field thereby reducing contrast immediately. In general, the disadvantages of these after-effects are often insufficient, in the writer's opinion, to offset the advantages of the higher intensity and good quality white light that may be obtained when suitable phosphors are used in the mercury-argon fluorescent lamps.

The switching circuit for exciting these lamps should provide virtually simultaneous extinction of one lamp and ignition of the other. This reduces flicker, a difficulty since the lamps (with or without phosphor) extinguish so rapidly. Kupperian and Golin accomplish this by a circuit in which the firing of one lamp extinguishes the other by means of a condenser which couples the anodes of both lamps.[5] Since the switching action of their circuit depends on firing the extinguished lamp, an additional circuit was necessary to provide a high-voltage surge to fire the second lamp. Merryman and Allen use a mechanical SPDT relay to switch the tubes, but make no provision for a higher voltage at the moment of switching.[6] This necessitates a power supply of much higher voltage than is actually needed to run the lamps since the higher voltage has to be available during switching.

The switching circuit shown in Fig. 3 provides the practically instantaneous transfer characteristic of the Kupperian-Golin circuit as well as a high-voltage surge at the moment of switching so that the higher striking voltage can be obtained from a lower voltage power supply. This can be done more simply than in the Kupperian-Golin circuit because the *extinction* of one lamp fires the other lamp. In addition, the present circuit uses the same circuit elements to provide an arrangement for controlling the duration of the light pulses. The present circuit thus has the following desirable features: (1) instantaneous switching; (2) high-voltage surge at the moment of switching; (3) integral timing of the light pulses; and (4) noiseless operation. The first two items are characteristic of the Kupperian-Golin circuit and none of them is characteristic of the Merryman-Allen circuit.

The high voltage surge in the present circuit is derived from a type of circuit often used to start fluorescent lamps operated on direct current. Essentially such a circuit consists of a switch, connected across the lamp, both of which are in series with an inductance. The voltage applied to the lamp is insufficient to start it. The switch in parallel with the lamp is closed momentarily and then opened. Opening the switch, S1, provides the voltage surge that builds up across the lamp because of the collapsing inductive field. In the present circuit the stimulus- and background-lamps are in parallel with one another and in series with an inductance. The extinction of one lamp, acting as does the switch described above, provides the high-voltage surge that fires the other lamp.[7]

---

[5] Kupperian and Golin, *op. cit.,* 275.
[6] Merryman and Allen, *op. cit.,* 174.
[7] The writer is indebted to Dr. John H. Greig for suggesting this application.

The timing and switching are accomplished by a multivibrator circuit making use of the screen and control grids of the two pentodes (V1 and V2). L1 is normally fired. The circuit is prevented from cycling by the thyratron (V3) that is normally biased, through S2, to cut-off.[8] A positive pulse is applied to the thyratron grid

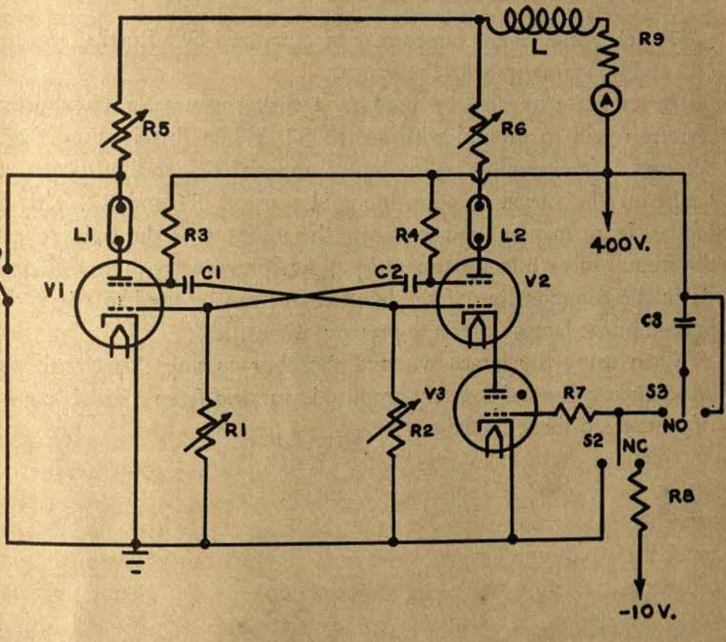


FIG. 3. COMBINATION TIMING AND SWITCHING CIRCUIT

L	= 12 h	A	= 0-50 milliammeter
R1, R2	= 3 megohm, variable	V1, V2	= 6V6
R3, R4	= 50,000 ohm	V3	= 2D21
R5, R6	= 10,000 ohm, variable	L1, L2	= fluorescent lamps, 1.2 x 4 in.
R7	= 1 megohm		
R8	= 100,000 ohm	S1	= SPST push-button switch, normally open
R9	= 20,000 ohm		
C1, C2	= 0.5 mfd., oil-filled	S2	= SPDT switch
C3	= 0.005 mfd.	S3	= SPDT lever switch

by operating S3. V2 and V3 start to conduct and C2 discharges through R1 thereby biasing V1 to cut-off, which extinguishes L1 and causes L2 to fire as explained above. After C2 has discharged, V1 again starts to conduct and causes C1 to discharge through R2. The negative voltage thus produced at the grid of V2 biases this

[8] This circuit can be modified to eliminate the thyratron by replacing C_1 with a resistor and biasing the grid of V2 to cut-off with a cathode resistor. (See H. J. Reich, *Theory and Application of Electron Tubes,* 1944, 359-360.) The thyratron circuit is preferred by the writer because the circuit values are not so critical and it is easier to equate the lamp intensities.

tube to cut-off, extinguishing L2, which then ignites L1. The thyratron, which is already biased to cut-off, prevents the action of the circuit from continuing and L1 continues to fire with L2 extinguished. Note that for a cold start it may be necessary to ignite L1 by actuating the push-button switch S1. Once L1 is ignited the switching action described above actuates the lamps.

The intensities of the lamps may be equalized by adjusting the resistors (R5 and R6) in series with the lamps.

The circuit may also be used as a flicker generator by shorting the thyratron grid to ground with switch S2. When this is done the circuit becomes a free-running multivibrator and the on- and off-times of both lamps may be varied by adjusting grid resistors R1 and R2. This feature of the circuit may be used to equate the intensity of the light reflected by the stimuli in each compartment by flicker photometry. The milliammeter, A, in the common line of the lamp circuit can be used to check the calibration of the lamps or reset to previous intensities.

When using the lamps without phosphor coating this circuit provides rectangular pulses of constant amplitude varying from a reliable minimum of 5 m.sec. to about 2 sec.

PHOTOELECTRIC RECORDING OF EYELID MOVEMENTS

By C. M. FRANKS and W. C. R. WITHERS, University of London

The simplest method of measuring eyelid movements—that of direct observation—is inaccurate, tedious, and provides little information. Many other methods—mechanical, electrical, and photographic—have been used,[1] but they have required cumbersome apparatus, or placed undue restriction on *S*'s movement, or necessitated the attachment of electrodes, paper eyelashes, or other artificial objects to *S*'s skin. Of these methods, the most convenient by far is the electromyographic one, its main difficulty being that the records of eyelid activity are sometimes obscured by facial muscle-potentials.

The present method was designed especially for use with mentally disturbed patients. It does not require the attachment of electrodes. It is accurate, inexpensive, and convenient. The apparatus is stable, requires almost no adjustment during operation, and can be made very insensitive to the activity of the facial musculature.

The method of recording is based on the principle that a photoelectric cell passes a current proportional to the intensity of incident illumination. The cell employed in the present apparatus is a caesium cathode Mullard Miniature Gas Filled No. 90CG, which is appropriately sensitive to artificial light emitted from standard gas-filled lamps. Ordinary room lighting can therefore be employed, provided that a high constant level is maintained, but the fixtures should be mounted on rigid metal supports. Unless the distribution of light in the room is fairly homogeneous, movements of *S*'s head will produce changes in the intensity of illumination falling upon the photoelectric cell. The more homogeneous the illumination, therefore, the less the restriction of movement required.

The photoelectric cell is mounted in a small screening can in one side of which a ½-in. hole is drilled to admit light. The can is mounted on a

[1] Raymond Dodge, A pendulum-photochronograph, *J. Exper. Psychol.*, 9, 1926, 155-161; A. L. Bernstein, Temporal factors in the formation of conditioned eyelid reactions in human subjects, *J. Gen. Psychol.*, 10, 1934, 173-197; H. E. King and C. C. Landis, A comparison of eyelid responses conditioned with reflex and voluntary reinforcement in normal individuals and in psychiatric patients, *J. Exper. Psychol.*, 33, 1944, 210-220; M. E. Bitterman, Heart rate and frequency of blinking as indices of visual efficiency, *J. Exper. Psychol.*, 35, 1945, 279-290; K. W. Spence and Janet Taylor, Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning, *J. Exper. Psychol.*, 42, 1951, 183-188; G. Gordon, Observations upon the movements of the eyelids, *Brit. J. Opthal.*, 35, 1951, 339-351; A. Carpenter, The rate of blinking during prolonged visual search, *J. Exper. Psychol.*, 38, 1948, 587-591.

plain perspex disk fitted in place of one lens in an ordinary spectacle frame (Fig. 1). The can is so small that *S*'s vision is comparatively unobscured. (If necessary, even smaller—but more expensive—photoelectric cells are obtainable.) When the eyelid moves, there is a change in the amount of light reflected on the cell. The varying *EMF* thus generated is fed into a linear amplifier by means of a screened lead. After amplification.

FIG. 1. THE PHOTOELECTRIC CELL MOUNTED ON THE SPECTACLES

this voltage is used to drive the pen of a recording milliameter (an Evershed and Vignoles model). In eyelid conditioning experiments, for which the apparatus was developed, the unconditioned stimulus (*US*), a puff of air, is delivered to the eye through a nozzle which is mounted in a small aperture cut in the perspex lens about ¾ in. below the cell.

The amplifier, shown diagrammatically in Fig. 2, consists of two stages —a directly coupled amplifying stage and a cathode coupled valve-volt meter circuit. To reduce hum as far as possible, the photocell load is also the grid leak for V_2, and is mounted inside the screened grid cap of this valve. Since the coupling is direct, a constant positive potential is applied to the grid of V_2; this arrangement necessitates a greater bias on V_2 than can be provided by R_{11} alone, without introducing a large amount of negative feedback and reducing the gain. To provide this bias, R_2 is used to pass additional current through R_{11}. V_2 is AC-coupled to V_3, that small movements of the head relative to the photocell or light source only temporarily deflect the recording meter from its zero-position. Initially, direct coupling was use, but it was abandoned because it required *E* to spend most of his time in resetting the recorder. The time-constant of the coupling (5.7 sec.) was so chosen as to give a fairly quick recovery, without undue distortion of the blink-record.

The valve-volt meter circuit (V_3) consists of a pair of cathode-coupled triodes with the recorder connected between anodes. F_2 is used to zero the meter by varying the current through one of the triodes. S_1 is a shorting switch which is used to short out the recorder while connecting up the *S*

or making other major adjustments. The valve-volt meter circuit is very stable since any changes in heater voltage (HT) affects both triodes equally, and leave the meter-current practically unchanged. R_8 is included that the grid current may be limited to a safe value. R_7 enables the maximal gain to be fixed at a suitable value under any reasonable lighting conditions, P_1 being the variable gain-control. The power-supplies are conventional. The photocell and V_2 are supplied from a 150-v. source, neonstabilized

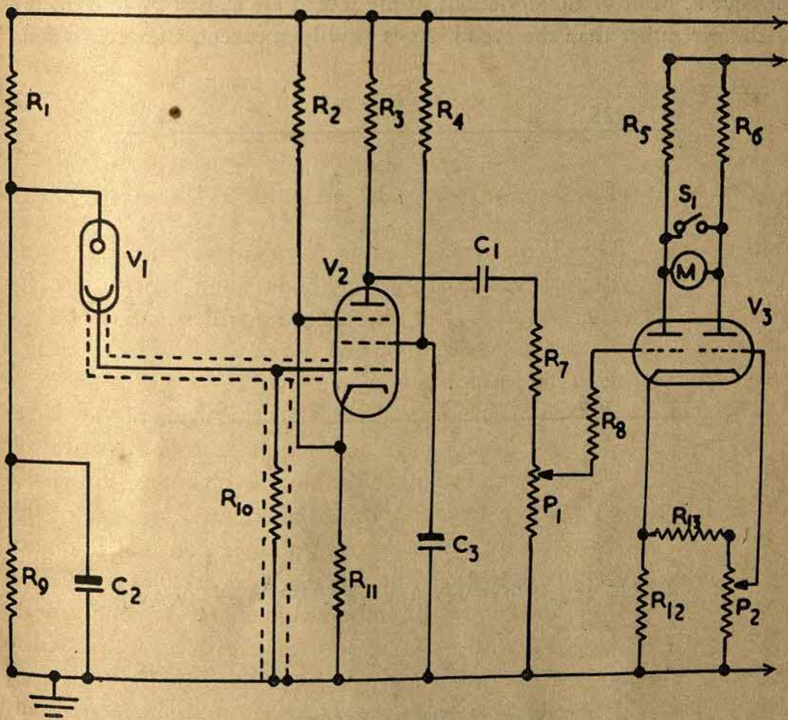


FIG. 2. WIRING DIAGRAM OF THE AMPLIFIER

R_1, 57,000 Ω	R_7, 4,700,000 Ω	R_{13}, 100,000 Ω	S_1, SPST
R_2, 33,000 Ω	R_8, 6,800,000 Ω	P_1, 1,000,000 Ω (linear)	M, recording milliameter
R_3, 220,000 Ω	R_9, 57,000 Ω	P_2, 50,000 Ω (linear)	V_1, photoelectric cell, type 90 CG
R_4, 570,000 Ω	R_{10}, 10,000,000 Ω	C_1, 1 microfarad	V_2, EF 37 A (6J7)
R_5, 22,000 Ω	R_{11}, 670 Ω	C_2, 8 microfarads (electrolytic)	V_3, 6SN7
R_6, 22,000 Ω	R_{12}, 1,000 Ω	C_3, 8 microfarads (electrolytic)	

to prevent changes in HT from appearing on the record. V_3 is supplied from a 250-v., unstabilized source.

As reported elsewhere,[2] the apparatus was very successful when used in studies of eyelid conditioning. Two typical records are presented in Fig. 3. In each of them the time-sequence reads from left to right and the upper tracing acts as an event-marker. The two events shown are the occurrence of the unconditioned stimulus (*US*)—a puff of air—and the conditioned stimulus (*CS*)—a tone—; the time-interval betweeen these two events being of the order of 25 sec.

The upper record is that of an *S* whose spontaneous blinking was very infrequent. Most of the deviations in the record are caused by movements of the eye rather than the eyelid. As is readily apparent, these deviations

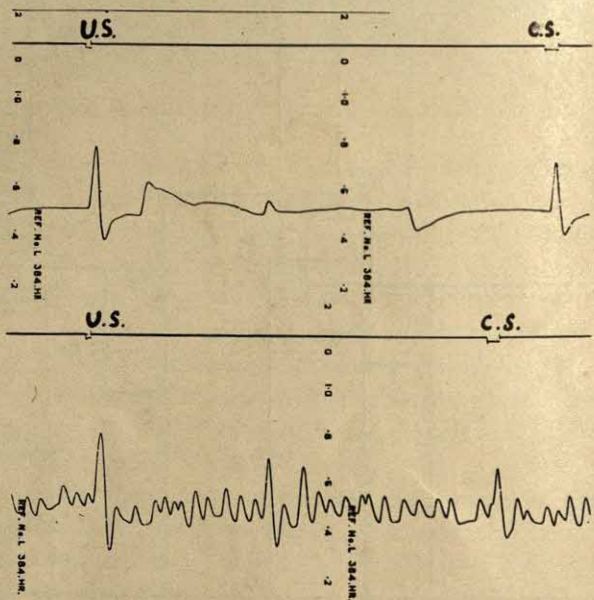


FIG. 3. SAMPLE RECORDS

bear little resemblance to the blinks produced by the *US* and the *CS*. The lower record is that of an *S* with considerable spontaneous blinking of varying amplitudes. The usual reference-grid printed on the recording chart has been omitted in Fig. 3 for reasons of clarity; when the actual chart is used, these markings make it very easy to measure the amplitudes, latencies, and deviations of all responses.

[2] C. M. Franks, An experimental study of conditioning as related to mental abnormality, Ph.D. Thesis, University of London Library, 1954, 90-103.

The apparatus is suitable even for *S*s who insist upon keeping their eyes closed throughout the conditioning procedure. A slight increase in the gain of the amplifier makes it possible to record the 'screwing up' movements and contractions of the closed eye produced both by *CS* and *US*. An adaptation of the apparatus which may be more conventient under certain conditions involves the use of a photoelectric cell which is maximally sensitive to infra-red light. Such an apparatus could be used in a darkened room.

APPARATUS NOTES

AN INEXPENSIVE AND ODORLESS RAT COLONY

One consideration tending to discourage animal research in small colleges is the expense involved in establishing a colony. Recently, the Department of Psychology of Carleton College set up such a colony at small cost. It is compact and relatively odorless, even though as many as 100 rats are housed in it along one wall of a small room.

The rats live in clear plastic cages, adapted from Admiral refrigerator vegetable hydrators (model 3698), which sell for less than $3 apiece. Each cage, excluding a sloping end, measures 9 × 13.5 × 7.5 in., and is large enough to accommodate three mature rats (see Fig. 1). Animals so housed have remained clean and healthy.

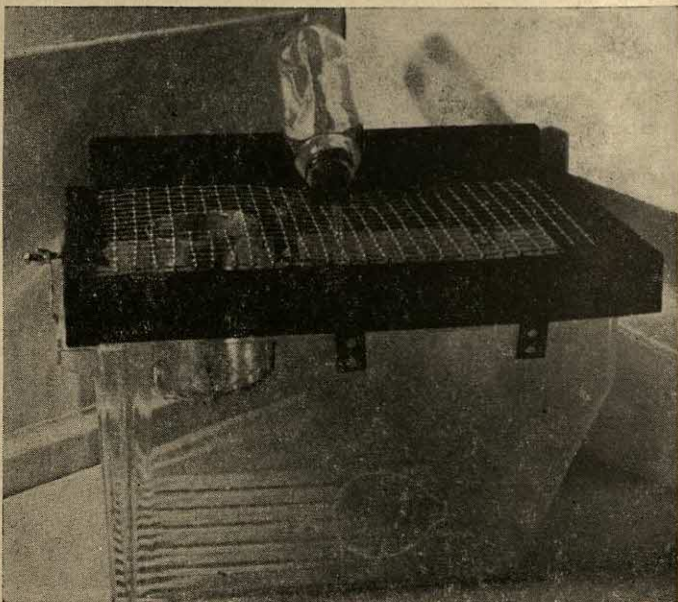


FIG. 1. ANIMAL CAGE

The cage cover consists of a wood frame with a hardware cloth top. This cover provides sufficient ventilation, and the additional height afforded by its sides increases the inside dimensions of the cage. Portions of the frame may be lined with metal to prevent gnawing, or as an alternative, the top may be made entirely of metal. The cover is anchored to the plastic cage by metal angles, which engage a lip protruding from the hydrator. The angles are bolted to the cover and may be disengaged by loosening a butterfly nut, affording easy removal of the cage top.

For feeding, a No. 2 tin can, with both ends removed, is suspended inside the

cage from the cover. The bottom of the can is replaced with wire mesh, through which the animal chews its food. A half-pint bottle, fitted with a single-hole rubber stopper and glass tubing (7 mm. outside diameter, bent at about 45°), serves as an adequate watering device. The bottle rests in a V-cut in the cover frame, and the tubing is inserted through the mesh top. By means of these devices, feeding and watering of the rats are accomplished without opening the cage, and the lightweight housings, since they are independent units, may be used as carrying receptacles when the animals are needed for experiments or classroom demonstrations.

The cage floor is covered with wood shavings which prevents odors from becoming objectionable if the cage is cleaned every few days; the frequency depending on the number and size of the rats housed. The cages may be washed easily and need not be dried before reuse. Ordinarily, only about one and one-half hours are required to clean, feed, and water 35 cages. The total amount of maintenance necessary is such that it can be accomplished by a single student working in his spare time, and represents little drain on a departmental budget.

The colony-room is ventilated by an exhaust fan attached to a window. With this ventilation and the cleaning schedule described above, it has been possible to maintain a sanitary and odorless colony.

Such a colony may be fitted into a restricted space. The cages are placed on three shelves, each 12-ft. in length, arranged one above the other, which suffices for 35 cages. Thus, only one wall of the room is occupied by the cages, and it is possible to utilize the remaining space for other equipment, such as a sink and washing board and activity cages.

There are three principal deterrents to the establishment of rat colonies: (1) they are too often expensive; (2) they may require too much space; and (3) they may become unsightly and malodorous. The writers feel that a colony such as has been described here meets all of these objections.

Carleton College

SUMNER HAYWARD
ROBERT ADAMSON

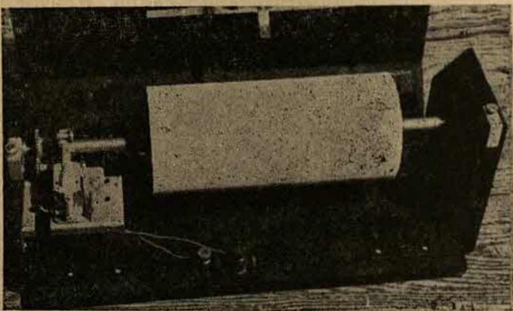
A DRUM FOR TACHISTOSCOPIC PRESENTATION

The apparatus to be described may be used as a memory-drum or for any other purpose demanding tachistoscopic presentation of verbal material. There are several objections to the drums currently available for such uses. They ordinarily are activated by either a friction drive or by lugs which serve as ratchets, which in turn are advanced by an arm protruding from the shaft of a constant speed motor. The friction drive design suffers from variation in the resiliency of spring washers which maintain the friction; further, this design usually incorporates a pulley system operating a rubber belt, which exhibits an unfortunate tendency to break. Most of the ratchet operated drums perform well within their limitations, if they are very carefully constructed. Even spacing of the ratchet lugs is difficult to obtain, but is essential to prevent a variation in the extent of the advance of the drum with each impulse. In addition, the number of advances per unit time of these drums is limited to one ratio.

To meet these objections, the writer has constructed a drum which combines the advantages of low cost construction (less than $10) and stable operation.

It is mounted on a half-inch shaft, the ends of which are fitted into ball bearings. The shaft need not be of this size, but it is a convenient one since it allows the use of conventional kymograph drums. The bearings are so seated in V-grooves that the drum may be lifted out of the apparatus for easy replacement of the verbal material to be presented.

The shaft is activated through a gear-train. One gear is pressed onto the shaft itself, this gear being operated in turn by a gear attached to the shaft of a ratchet-



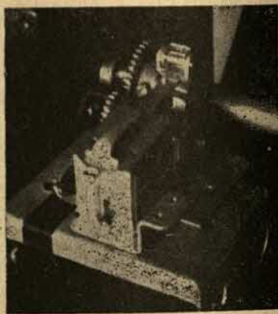


FIG. 1. THE DRUM WITH COVER RAISED AS VIEWED FROM THE REAR

FIG. 2. GEAR-TRAIN AND RATCHET-SOLENOID

solenoid. Several such solenoids are available commercially. The one used by the writer is the Guardian Locking Relay, Type RC—100—AR, which sells for slightly over $5. This relay advances a twelve-tooth ratchet gear which is mounted on a quarter-inch shaft. The drive-gear may be pressed onto this shaft to mesh with the gear on the drum-shaft. The drum-gear should be a multiple of the drive-gear in order to insure exact spacing of the drum-travel with repeated revolutions. Thus, the number of advances for each revolution of thd drum may be any convenient multiple of twelve (or the number of teeth of the ratchet gear). The present model utilizes a solenoid-to-drum ratio of one to two, which yields 24 impulses for each drum revolution.

The drum may be of any size, limited only by weight, which, if too great, would provide too much inertia for a rapid succession of impulses. The limiting number of impulses for the solenoid is about 10 per second. The solenoid is activated by an impulse timer, which is commercially available at small cost, or which may be easily constructed.

Carleton College ROBERT ADAMSON

AN IMPROVED DEVICE FOR RECORDING CHANGES IN SKIN-TEMPERATURE

Devices for recording changes in skin-temperature are of interest to many investigators. Basic to instruments of this kind is a highly sensitive pick-up for registering temperature. Such a device was the thermistor described by Baker and Taylor in

1953.[1] Since the thermistor has since been improved, new and better devices are now available.

The thermistors used in the earlier device had to be welded into the recording circuit.[2] This task was tedious and often resulted in ruined thermistors. In addition there was the desire for greater sensitivity and the problems of controlling lag, drift, air-temperature changes, and of avoiding the possible influence of the galvanic skin response (*GSR*). The device described here (Fig. 1) possesses greater sensitivity and it escapes or greatly reduces many of the problems mentioned.

The thermistor used in this pick-up is a glass-covered bead made by Western Electric (Code #D177232). The device utilizes a bronze spring clamp which provides a light, constant contact that does not interfere to an appreciable extent

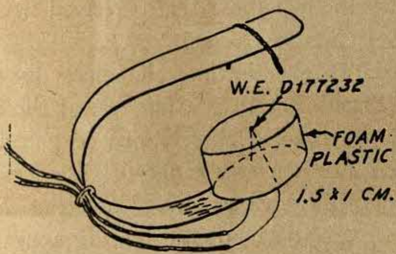


FIG. 1. A PICK-UP ELECTRODE DESIGNED FOR USE ON THE FINGER TIP

with blood-flow in the capillaries. The device is used on the finger tip which is subject to some temperature change as circulation of blood in the capillaries varies under stress conditions. There is also a rather satisfactory avoidance of difficulty in recording due to movement about the contact point. There is no risk that *GSR* can be picked up because of the glass covering. Slight modifications would enable *E* to obtain readings from other areas of the body.

The thermistor is imbedded in sponge plastic which shuts out the effects of random air currents in the room and reduces temperature changes due to radiation of heat from surrounding objects. The sponge has low heat absorption which is desired because of drift and lag associated with mass of most materials.

In the use of thermistors to measure variations of skin-temperature, it is desirable to employ a bridge and amplifier circuit of good stability, and to regulate the voltage supplied to it. For very accurate research work careful control of room temperature and humidity is considered essential. The object in such recording is to obtain variations arising from internal bodily changes and to avoid all temperature changes having an external source.

Purdue University

LAWRENCE M. BAKER
GEZA A. CSAPO

[1] L. M. Baker, and W. M. Taylor, An apparatus for recording changes in skin-temperature, this JOURNAL, 66, 1953, 124-125.

[2] Baker and Taylor, The relationship under stress between changes in skin temperature, electric skin resistance, and pulse rate, *J. Exper. Psychol.*, 48, 1954, 361-366.

NOTES AND DISCUSSIONS

THE RELATION OF PERCEIVED SIZE TO PERCEIVED DISTANCE: AN ANALYSIS OF GRUBER'S DATA

In a recent paper, Gruber presents data on perceptions of size and distance which he takes to indicate that perceived size is *not* related to perceived distance.[1] Actually, he has obtained some evidence to support the hypothesis he rejects—the hypothesis that perceived size is proportional to perceived distance.[2]

Gruber performed two experiments. He rejects Experiment I for this purpose and offers Experiment II as a fair test of the hypothesis. In this experiment he obtained coupled observations of perceived size and perceived distance from 20 *O*s over a range of distances from 200 to 450 cm. Within this range *O* viewed a standard object at each of six different distances and so set a variable object that it appeared half-as-distant (½D); equally distant (ED): equal in size (ES½) with the standard half as far from *O* as the variable; and equal in size (ESE) with the standard and variable equidistant from *O*. Although unequivocal instructions are not reported, Gruber states that he instructed his *O*s to adopt a 'phenomenal' as opposed to an 'objective' attitude.[3]

The purpose of the present note is not to criticize Gruber's experimental procedure nor to question the validity of his results. Accepting his findings at face value, the question is simply: Do these results in fact reveal a situation in which relative size is *under*estimated when relative distance is *over*estimated—a size-distance paradox—as Gruber asserts? An alternative interpretation may better fit the facts and remove the paradox into which Gruber's discussion has plunged us.

Let us reëxamine the mean size and distance judgments shown in Table III of Gruber's paper. These data are reproduced here (Table I) together with the corresponding values of perceived size and perceived distance obtained by performing some very simple calculations on the recorded data.

---

[1] H. E. Gruber, The relation of perceived size to perceived distance, this JOURNAL, 67, 1954, 411-426.

[2] A. S. Gilinsky, Perceived size and distance in visual space, *Psychol. Rev.*, 58, 1951, 460-482.

[3] For a recent study of the importance of instructions see Gilinsky, The effect of attitude upon the perception of size, this JOURNAL, 68, 1955, 173-192.

The first and second columns of Table I give the physical distances D_1 and D_2 of the variable and the standard stimulus-objects respectively. The third column gives the physical distance D separating the variable and the standard. Subtracting the settings of distance judged half (½D) from the corresponding settings of distance judged equal (ED) gives six measures of perceived distance *d*. The obtained size-judgments may be conveniently converted into six corresponding measures of perceived size by dividing the matched size of the standard (ESE) at distance D_1 by the matched size of the standard (ES½) at distance D_2. These ratios of perceived sizes, s/S,[4] may be plotted directly against the corresponding

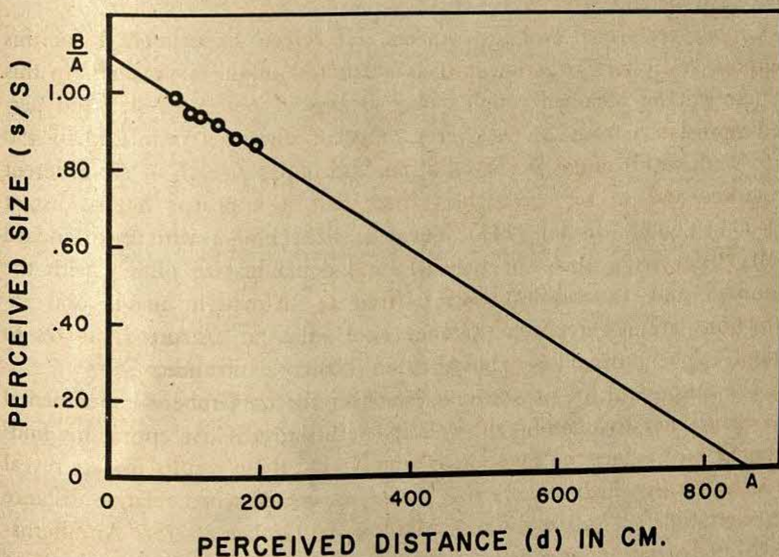


FIG. 1. THE RELATION OF PERCEIVED SIZE TO PERCEIVED DISTANCE
A straight line has been fitted by eye to the data obtained by Gruber. The intercepts, B/A on the vertical axis, and A on the horizontal axis, give the two parameters, A and B, directly from the graph.

values of perceived distance, *d*. The resulting graph (Fig. 1) is a straight line, inclined downward to the right. The relation is given by

$$s/S = (B/A)[(A\text{-}d)/A] \quad \ldots\ldots\ldots\ldots\ldots \quad [1]^{5}$$

[4] This notation will be used henceforth as being both simpler and more general in meaning than that used by Gruber. For definitions and discussion of these and other symbols used in this paper see Gilinsky, *op. cit., Psychol. Rev.*, 58, 1951, 462-469.

[5] Gilinsky, *op. cit.*, 482.

For d = 0, s/S = B/A. For s/S = 0, d = A. Accordingly the straight line graph intercepts the vertical axis at B/A and the horizontal axis at A. The two parameters, A and B, are thus given directly by the plotted graph. A = 854 cm. and B = 930 cm.

The values of the parameters thus obtained by extrapolation of the straight line of Fig. 1 may be confirmed by checking against the remaining data. The results thus far analyzed are independent of physical distance, that is to say, the physical distances used in the experiment did not enter the calculations on which the numerical values of A and B are based. However, we should be able to predict the relation between perceived size and physical distance D by substituting the numerical values of A and B in the basic equation for perceived size derived in a previous paper.[6] This equation is

TABLE I

OBTAINED AND THEORETICAL VALUES OF PERCEIVED SIZE AND DISTANCE

Data (in cm.) from Gruber's Table III.

Physical Distance			Perceived Distance				Perceived Size			
			obtained			theo.	obtained			theo.
D_1	D_2	D_1-D_2	ED	½D	d	d	ESE	½ES	s/S	s/S
450	225	225	449	260	189	178	10.6	12.3	.862	.862
400	200	200	400	238	162	162	10.3	11.7	.880	.882
350	175	175	347	206	141	145	10.5	11.5	.913	.904
300	150	150	296	179	117	128	10.3	11.1	.928	.926
250	125	125	248	144	104	109	10.4	11.0	.945	.950
200	100	100	199	115	84	90	10.2	10.4	.981	.975

$$s/S = B/(A + D) \qquad [2]$$

For each coupled observation of perceived size and perceived distance the physical distance D is known. Under the conditions of Gruber's experiment where the distances of the standard and variable stimulus-objects were varied together to maintain a constant ratio, the distances (D_1-D_2) separating the standard and variable are equal to the distances (D_2) of the standard from *O* and serve as the measures of D. By substituting the values of A and B obtained directly from Fig. 1 we may calculate the theoretical ratios of s/S for comparison with the obtained ratios of s/S of reach recorded value of D. Note the correspondence of the paired values in Columns 10 and 11 of Table I. Fig. 2 is the graph

[6] *Ibid.*, 467.

of perceived size as a function of physical distance. The curve drawn through the data points is that of Equation [2] using B = 930 cm. and A = 854 cm. as determined above.

The close correspondence between the theoretical and the obtained values of perceived distance (d) may be similarly shown (Table I) by

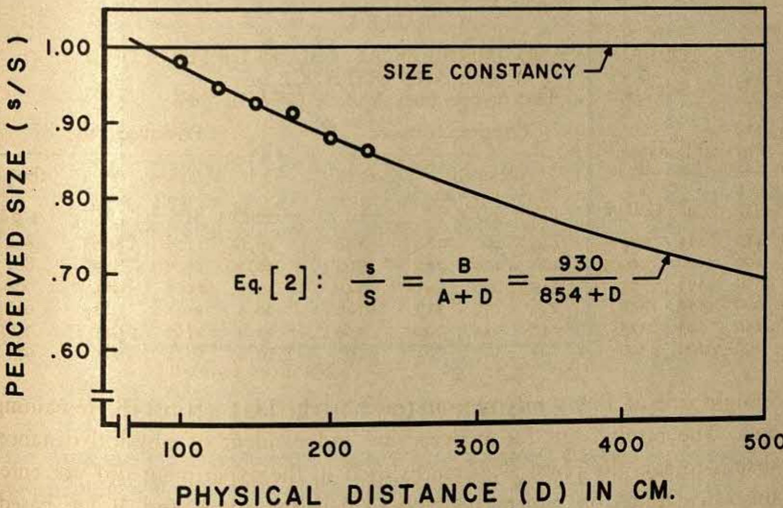


FIG. 2. PERCEIVED SIZE AS A FUNCTION OF PHYSICAL DISTANCE
The data from Gruber are fitted by the curve of Equation [2], using the values of the parameters, A = 854 cm., and B = 930 cm. obtained from Fig. 1.

means of the associated equation for perceived distance as a function of physical distance.[7] This relation is

$$d/D = A/(A + D) \quad \cdots\cdots\cdots [3]$$

Fig. 3 is the graph of this function using the identical value of the parameter A (A = 854 cm.). Note that this formulation and the resulting graph do not take measurements of size into consideration.

The three variables, perceived size, perceived distance, and physical distance are functionally interdependent. Given the values of any two of these variables, we are able to predict the values of the third variable. Gruber asserts that his results represent "another series of failures to find any relation between perceived size and perceived distance."[8] The present

[7] *Ibid.*, 466.
[8] Gruber, *op. cit.*, 420.

analysis finds that his results show a definite and mathematically predictable relation between perceived size and perceived distance and that from this relation the parameters which govern both size and distance

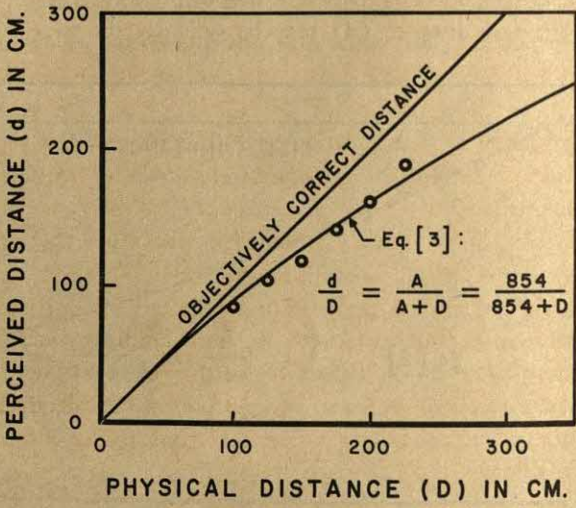


FIG. 3. PERCEIVED DISTANCE AS A FUNCTION OF PHYSICAL DISTANCE
The data from Guber are fitted by Equation [3] using only the value of the parameter A (854 cm.) obtained from Fig. 1.

perception may be empirically determined. The use of Gruber's data for the above demonstration will, it is hoped, dispel the atmosphere of paradox from his findings and reveal their basic significance.

Columbia University ALBERTA S. GILINSKY

THE OPTICAL EXPANSION-PATTERN IN AERIAL LOCOMOTION

Perceptual-motor tasks, especially tracking and pursuit tasks, have been much investigated by psychologists during and since World War II. The problem of designing control systems for use by human operators and the need for considering the overall 'man-machine' system have made this study imperative.[1] The particular tasks or acts studied, however, have

* This research was supported in part by the United States Air Force under Contract No. AF330(038)-22373, monitored by the Skill Components Research Laboratory, Air Force Personnel and Training Research Center. The writer is indebted to Dr. D. M. Purdy for his assistance in this study.

[1] P. M. Fitts, Engineering psychology and equipment design, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 1287-1335.

generally been ones which could conveniently be set up in a psychological laboratory, they have tended to be manipulative, and there has been little or no research on the fundamental act of *locomotion*. It might be argued that the performance of locomotion is more important than the performance of manipulation, since many more animals are capable of the former than the latter. Some psychologists, notably Lewin, have taken locomotion to be the prototype of all behavior.[2] The fact that man can perform locomotion by manipulation when he uses powered vehicles such as the airplane should not make us lose sight of the primary performance of 'going somewhere.'

Locomotor behavior is intimately bound up with visual space-perception. Whether an animal walks or runs on the ground or, like a bird-flies through the air and returns periodically to the terrain, his behavior must be nicely adjusted to the surface of the earth and also to the complex of surfaces that clutter up the face of the earth. The governing or guidance of this behavior depends in large part on optical stimulation.[3] The flow of optical stimulation provides a 'continuous feedback' of information for controlling the flow of the motor activity.

Among other principles, changing optical stimulation is governed by the principle of motion parallax or, more generally, motion-perspective.[4] Motion in the retinal image provides at once information about space and about the movement of its possessor. The writer and collaborators have recently presented a mathematical description of motion-perspective, in which two 'patterns' of optical motion were defined.[5] The case considered was the abstract one of an eye moving in any direction relative to a textured plane surface, and the rays of light reflected from that surface.

A frontal surface perpendicular to the line of locomotion always yields a radial pattern of velocities, whereas a longitudinal surface parallel to the line of locomotion always yields a 'unidirectional' pattern of velocities. The first might be called 'pure expansion' and the second 'pure flow.'

---

[2] Kurt Lewin, *Principles of Topological Psychology*, 1936; *passim*.

[3] A possible exception is the behavior of bats and of the blind, whose locomotion seems to be controlled by auditory stimulation in the form of echoes. See D. R. Griffin, Sensory physiology and the orientation of animals, *Amer. Scientist*, 41, 1953, 209-224; and the studies on "facial vision" by K. M. Dallenbach and his students, this JOURNAL, 57, 1944, 133-183; 60, 1947, 502-553; 63, 1950, 485-515; 66, 1953, 519-553.

[4] J. J. Gibson, *The Perception of the Visual World*, 1951, 117-136; K. M. Dallenbach, The elastic effect: An optical illusion of expansion, this JOURNAL, 66, 1953, 634-636.

[5] J. J. Gibson, Paul Olum, and Frank Rosenblatt, Parallax and perspective in aircraft landings, this JOURNAL, 68, 1955, 372-385.

They can be plotted on angular coördinates to give a kind of visual picture of the stimulus in either case. For the human flier, the first is exemplified by the case of a vertical helicopter landing and the second by the case of straight and level flight. A glide corresponds to a sort of mixture or fusion of the two optical patterns.

These modes of stimulation apply in cases other than human flight. The expansion pattern exists during a fall toward the ground under the influence of gravity, for instance, or during an automobile driver's approach to a brick wall, or when a man gets down on his knees to examine a blade of grass. The flow pattern exists for the surface of the earth during walking, running, riding, or driving any wheeled vehicle.

In general, then, pure expansion in the field of view is a sensory symptom of approach. It implies future contact with something. Whenever contact of any sort is the goal or motive of locomotor behavior, it might be assumed that pure expansion would be a 'desirable' sensory impression. On the other hand, contact may involve collision. This event is painful and is presumably not a goal or motive of locomotion. Hence it might also be assumed that pure expansion would be an 'undesirable' sensory impression.

Pure flow in the field of view is a sensory symptom of progression. In an eloquent idiom it is a sign of 'covering' ground. As every automobile driver knows, the faster the flow the more ground he is covering. So long as there is no component of pure radial expansion in the field of view the individual who moves is justified in feeling 'safe.' The flow of the ground contributes information that one is going somewhere and the direction of flow tells one where, but it is not a symptom either of 'arriving' or of potentially 'bumping into' something.

The cessation of flow or expansion in the field of view is an indicator of coming to rest, just as much as the continuation of motion-perspective is an indicator of one's movement. Presumably this kind of stimulus-change might be a guide to the checking or decelerating of locomotion. A static field of view is an optically exact indicator of a stationary position.

The focus of flow or expansion in the visual field is an indicator of the direction of locomotion. Whatever object or bit of the terrain is at this optical focus at any moment is the part of the world toward which the eye is moving at that moment. The rule holds whether the spot is on the horizon of the surface at a very great distance, or is a spot of a surface which is being approached on a 'collision course.'

The continuous and symmetrical expansion of a bounded shape in the

field of view until it reaches a critical rate, or until its size extends over a critical area of the field is an indicator of imminent collision.[6] The continuous expansion of a 'texture' of the sort reflected by the elements of a surface is an indicator of the same thing.

In view of these sensory indicators, symptoms, or cues, what is implied for the governing and control of locomotion—for locomotion considered as the motivated or 'voluntary' act of an individual?

Application to aerial locomotion. Some of the special tasks involved in locomotor behavior might be designated as the 'aiming' of locomotion, the 'checking' of locomotion, the 'avoiding' of obstacles, and the 'pursuing' of a moving object. The landing of an airplane is a special case of aiming and avoiding in which, since the checking of speed is largely impossible, the task of levelling off is substituted.

In all these special tasks the optical pattern of flow or expansion—motion-perspective—is present to the eyes of the individual. The following hypotheses are offered as formulae or rules to account for the choice of motor-responses from moment to moment, in the case of aerial locomotion.

(1) Control of a landing glide. Those reactions should occur which keep the optical focus of expansion at the point of the ground which the pilot wants to get close to—usually the nearer end of the runway.

(2) Levelling off. As the stimulus for imminent collision grows more intense, such reactions should occur as to move the optical focus of expansion from the ground to the horizon, *i.e.* to shift from an expansion-pattern to a flow-pattern, and to complete this shift at the moment when the landing wheels are barely in contact with the ground.

(3) Helicopter landing. As the stimulus for imminent collision grows more intense, such reactions should occur as to cancel the expansion-pattern (to produce a static field of view) at just the moment when the landing gear is in contact with the ground.

(4) Avoiding obstacles. Such reactions should occur which keep the optical focus of expansion or flow away from expanding shapes in the field of view (objects) and within the area of the field *between* expanding shapes (ground or air).

(5) Pursuing another craft. Such reactions should occur as to keep the

[6] A current experiment by the writer has made the finding that an expanding shadow on a translucent screen viewed through an aperture by an *O* will induce an involuntary response of shrinking in the *O*. One condition of this response seems to be that the shadow must expand to fill the translucent screen.

optical focus of expansion of the background (clouds or terrain) in coincidence with the evasive shape of the fugitive object.

It is interesting to note that these locomotor tasks are interestingly related to the so-called 'pursuit tasks' which have been devised to study the human operation of control mechanisms.[7] At least one of the above tasks (5) is analogous to 'following pursuit' and another (1) is analogous to 'null pursuit,' but there is a fundamental difference between the two kinds of tasks: in locomotor behavior the entity under *O*'s control is not an object or an apparatus but is a visible point in optical space—a point produced by the act of locomotion itself. It is not an object at all but a kind of center of motion-stimulation.

The formulae given above are based on the assumption that locomotion, like other forms of motor action, is behavior with an end or purpose. The end is a state of affairs 'in space' which the individual is set to achieve. One such spatial condition is the reaching of a stationary destination—a place or an object. If the object itself moves, it is a 'fugitive destination' requiring pursuit. The avoiding of collisions is a subordinate spatial condition to be achieved.[8] The discrepancy at any time between the desired spatial condition and the present spatial condition is the stimulus for a reaction. Control reactions are such as to cancel the discrepancies, that is, they are compensatory in nature and cease when stimulation reaches a sort of equilibrium state. The fact that 'servo systems' can be designed to function in much the same way suggests that the process is less psychic than might be supposed.[9]

Cornell University J. J. Gibson

THE USE OF FOREIGN LANGUAGES BY PSYCHOLOGISTS

There are at least two objectives for the publication of scientific literature: (1) the current interchange of data and ideas among interested scientists; and (2) the permanent recording of these data and ideas. To achieve the first of these in international scientific communication requires the use of literature published in languages and countries other then that of the author. The references cited by authors might be expected to reveal the extent to which communication takes place internationally.

---

[7] Fitts, *op. cit.*, 1319-1331.

[8] J. J. Gibson and L. E. Crooks, A theoretical field-analysis of automobile-driving, this Journal, 51, 1938, 453-471.

[9] Fitts, *op. cit.*, 1327 ff.

The universe of literature to which scientists can refer is the sum of papers appearing in all countries and in all languages in which literature of the subject appears. If there is an adequate international communication in science then a writer in any country could be expected to draw from other national literatures in proportion to the distribution of materials by language and country. An hypothesis that the proportion of citations from different languages should be similar regardless of the nationality of the author can be tested by analyzing the references actually made in published papers. The present data are based on a distribution by language of all

TABLE I

LANGUAGE OF CITATIONS IN PSYCHOLOGICAL JOURNALS

Language	No. articles	No. refer.	Percentage of citations in				
			Eng.	Ger.	Fr.	Other	Own language
English (U.S.)	74	534	88.5	7.5	1.1	2.6	83.9*
English (U.K.)	22	345	98.0	1.7	—	—	30.7*
German	14	112	2.7	91.0	6.3	—	91.0
French	30	216	24.0	7.4	65.4	3.2	65.4
Italian	21	314	27.4	23.6	22.6	26.4	25.2
Spanish	17	65	47.5	20.0	6.2	26.3	24.5
Japanese	19	288	74.5	11.4	—	14.1	14.1
Total references		1874	64.0	15.0	12.0	9.0	
Psychological Abstracts (1952)		7297	77.0	5.5	8.0	9.0	

* References to publications of own country.

citations given by authors of papers appearing in the 1952 volumes of the following journals: *American Journal of Psychology, British Journal of Psychology, Schweizscher Zeitschrift für Psychologie* (German and French), *Journal de Psychologie Normale et Pathologique, Revista de Psychologia General y Aplicada, Archivio di Psicologia, Neurologia e Psichiatria,* and *Japanese Journal of Psychology.* In these seven volumes there were 197 papers with 1874 citations to literature; the details by language are shown in the first two columns of Table I.

It is impossible to define exactly the universe from which citations might be drawn, although for the present purpose two may be described. The first is the total number of references cited. Of the 1874 references 64% were in English, 15% in German, 12% in French, and 9% in other languages. For the second the entries in the 1952 volume of *Psychological Abstracts* might be taken to indicate the relative frequency of publications

in different languages. Of the 7,000-odd entries in this volume 77% were in English, 6% in German, 8% in French, and 9% in other languages. It is evident that these two distributions agree only in the percentage of papers in 'other' languages. While the entries in *Psychological Abstracts* have a date range of only two or three years the proportions are similar to those reported by Fernberger at intervals since the late 1930s.[1]

The data of Table I give no support to the hypothesis of proportional citation. American writers cite English language papers in excess of the percentage expected; English writers almost completely limited themselves to their language. It is also of interest to note that data not included in Table I show American authors cite British publications to the extent of 3.7%, while British authors make such citations 30.7% of the time. Germans also almost exclude literature except their own. French authors refer to their own language far in excess of its relative proportion in the comparison groups, but here considerable attention is paid to the English literature. In the other three languages English is used from one quarter to three quarters of the time. In large measure this reflects the lack of literature in these languages. For example, in the 1952 *Psychological Abstracts* 1.4% of the entries were in Italian and the same percentage in Spanish, with less than ½% in Japanese.

Language barriers and lack of foreign publications account for these data in some measure, but other factors are of equal importance. It is not the purpose here to analyze the reasons or to suggest solutions. The hypothesis that scientists currently draw on the world's literature in their work is not supported. International scientific communication is not as free and open to the rank and file of scientific workers as is often claimed.

Wayne University C. M. LOUTTIT

A COMPARISON OF THE METHODS OF OLFACTORY STIMULATION: BLASTING VS. SNIFFING

Although olfactory thresholds obtained by sniffing appear to be reasonably reliable,[1] it is not at all certain that they mean the same thing as thresholds obtained by techniques which do not depend upon active inhalation by *O*. Advantage was taken, therefore, of the opportunity offered

[1] S. W. Fernberger, On the number of articles of psychological interest published in the different languages, this JOURNAL, 28, 1917, 141-150; 37, 1926, 578-581; 48, 1936, 680-684; 59, 1946, 284-290; A national analysis of the psychological articles published in 1939, *ibid.*, 53, 1940, 295-297.

[1] F. N. Jones, The reliability of olfactory thresholds obtained by sniffing, THIS JOURNAL, 68, 1955, 289-290.

by the collection of thresholds by the controlled-blast method for another purpose to check sniffing thresholds for three substances: safrol, n-butyric acid, and n-butanol.[2] Although not definitive, the results suggest some questions for research.

Data were collected for 45 *O*s. The controlled-blast thresholds were obtained by means of the apparatus described elsewhere.[3] Because the purpose of the main experiment was to obtain data for as many *O*s as possible in a reasonable time, one ascending series was used for each threshold. Sniff bottles were prepared by successive dilution with mineral oil, as in the sniff reliability work. Every *O* was checked on the sniff series immediately after his completion of the 20 blast thresholds. Recognition-thresholds were used throughout.

For n-butanol and safrol the median thresholds obtained by the two methods were very similar, with sniffing revealing slightly greater sensitivity. The respective medians (in terms of molar ratio to air) are: n-butanol, sniff 3.6×10^{-5}, blast 9.5×10^{-5}; safrol, sniff 1.2×10^{-8}, blast 7.3×10^{-7}; n-Butyric acid, however, showed very different results. The medians are sniff 1.4×10^{-10}, blast 1.0×10^{-6}.

These results offer obvious difficulties of interpretation. The two methods check very well for n-butanol and safrol, but not at all well for butyric acid. Since a separate group of 24 *O*s sniffing a fresh series of dilutions gave very much the same result, I am not inclined to ascribe this divergent result simply to experimental error.[4] The most reasonable explanation, assuming the result is correct, is that the rate of adsorption of butyric acid is slow, or that it is very tightly bound once adsorbed, thus making the threshold very sensitive to differences in stimulus-volume. The blast stimulus consisted of only 40 ml. of odorous air, while the sniffs, although not measured, must have had much greater volume than this. This explanation is compatible with the probability, expressed elsewhere, that the most important process in olfactory stimulation is adsorption.[5]

---

[2] Jones, The analysis of individual differences in olfactory sensitivity, *Amer. Psychol.*, 9, 1954, 400-401 (Abstract).

[3] Jones, A test of the validity of the Elsberg method of olfactometry, this JOURNAL, 66, 1953, 82. See also, C. A. Elsberg and I. Levy, The sense of smell: I. A new and simple method of quantitative olfactometry, *Bull. Neurol. Inst. N. Y.*, 4, 1935, 5-19.

[4] It is interesting to note that this sniff threshold agrees very well with that obtained by a successive dilution method by Passy. His result expressed as molar ratio is 2.5×10^{-10}. See J. Passy, L'odeur dans la série grasse, *C. R. Soc. Biol.*, 45, 1893, 479-481.

[5] Jones, Olfactory absolute thresholds and their implications for the nature of the receptor process, *J. Psychol.*, (in press).

It would appear that more methodological work needs to be done before data collected by blast and sniffing can be considered to have the same meaning. If we raise the question of validity, as Blackwell has done for vision,[6] we must balance the 'background' odor often experienced from pure air during blast determinations against the obvious lack of stimulus-control in the sniffing technique. Probably if one's goal is the comparison of thresholds for the same substance under, say, different conditions of adaptation, the sniffing technique provides a rapid and simple method.[7] When one wishes to compare thresholds for different substances under rigidly controlled conditions so that the underlying physical and biochemical properties can be assessed, the controlled blast, using recognition thresholds to minimize artifacts, seems to be the method of choice.

University of California F. Nowell Jones
Los Angeles, California

CURVES OF INSULIN TOLERANCE FOLLOWING SUB-SHOCK DOSES OF INSULIN

While conducting a series of studies on the effects of sub-shock insulin on psychiatric patients, the writer found it necessary to determine the course of the hypoglycemia following the injection of insulin until the termination of the reaction.[1] This information made possible a study of the relationships between relative blood-sugar levels and behavior at given times.

Eighteen men, white, non-psychotic, neuropsychiatric patients who were receiving sub-shock insulin therapy served as *S*s. A fasting blood sample was taken from each *S* just before insulin was injected sub-cutaneously at 8:00 A.M. Six *S*s received 20 USP units regular insulin, six received 40 units, and six received 60 units. Subsequent blood samples were drawn at 15-min. intervals following the insulin injection until the reaction was terminated. The blood-sugar level of each sample was determined by the Folin-Wu method.[2] The importance of a possible potassium deficiency and its relationship to carbohydrate metabolism was dis-

[6] H. R. Blackwell, Psychophysical thresholds, *Engin. Res. Bull.* No. 36, University of Michigan, 1953, 1-227.

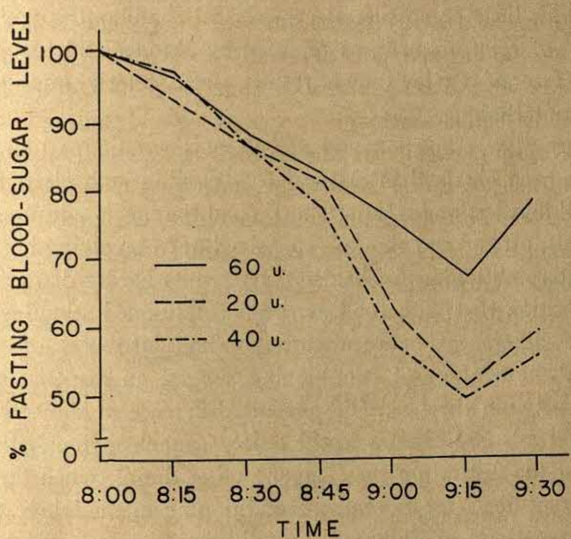
[7] G. H. Cheesman and Stella Mayne, The influence of adaptation on absolute threshold measurements for olfactory stimuli. *Quart. J. Exper. Psychol.*, 5, 1953, 22-30.

[1] P. R. Fuller, Some physiological and psychological effects of sub-shock doses of insulin. Unpublished Ph.D. thesis, Indiana University, 1952.

[2] O. Folin and H. Wu, Blood measurement, *J. Biol. Chem.*, 41, 1920, 367-374.

missed as a possible extraneous variable, when 12 electrocardiograms made on the same *S*s revealed no ischemia or hypokalamia.[3]

The insulin tolerance curves (Fig. 1) showed a decrease in blood-sugar level which began within 15 min. after injection and continued until a



minimal level was reached between 60 and 75 min. after injection. With one exception, *S*s' blood-sugar level began to return to normal before termination procedures were instituted by the staff.

Florida State University PAUL R. FULLER

OPERANT VS. CLASSICAL CONDITIONING

In the theoretical controversy between those who consider operant conditioning a *sui generis* form of learning and those who align it with classical conditioning, the first view seems to be gaining the upper hand. There is, however, one important *empirical* distinction between operant and classical conditioning that seems unknown—or at least unmentioned in the literature on *CR;* namely, the vast quantitative differences between the two with respect to extinction and partial reinforcement. In a recent review of more than 100 extinction experiments from Pavlov's laboratories I found that the extinction-scores of both the secretory and motor phases of their food conditioning seldom exceed a few dozen nonreinforced

[3] S. M. Perry and S. L. Rosenbaum, Hypopotassemia in untreated diabetic coma, *New England J. Med.*, 245, 1951, 847.

trials.[1] In operant conditioning, however, such scores run, as is well known, into hundreds and even thousands of trials.[2] Again, while effective classical conditioning has been produced by the Russians with reinforcements every other or even every third trial, lower ratios of reinforcements have met with little or no success.[3] Successful conditioning with a 1 : 50 'fixed ratio' of reinforcement, as reported by Skinner and his associates,[4] would, in face of Pavlov's own results, seem utterly impossible with Pavlov's own technique.

True, very slow extinction and successful low-ratio partial reinforcements have been obtained in classical conditioning with electric shock as the unconditioned stimulus.[5] But shock conditioning is admittedly a *CR*-class by itself, often—and rightly—conceived of as involving both classical and operant conditioning mechanisms. Its results hardly detract from the basic proposition that with food as the unconditioned stimulus the wide quantitative differences between corresponding attributes of a Pavlov-type and a Skinner-type of conditioning suggest an essential qualitative distinction between the two. The phyletic difference in subjects—dogs in one case and rats and pigeons in the other—cannot obviously be taken as the cause of the distinction since, among other things, when the Russians use an operant-like 'free-movement' conditioning methodology their findings approach those of Skinner.[6] The distinction or, we might say, the greater efficacy of operant conditioning inheres by all signs in the very essence of its type of modification. Moreover, for the present argument it matters but little whether this distinction or greater efficacy is conceived of as being due to a greater involvement of relational perception in operant conditioning or to some unspecified intraorganic behavioral mechanism that is somehow operative when 'conditioned stimuli produce conditioned responses and not when they merely accompany them.' What is important is the fact that such a distinction exists.

Queens College, New York GREGORY RAZRAN

[1] I. P. Pavlov, [*Pavlov's Wednesday Seminars*], 3 vols., 1949; [*Collected Works*], 5 vols., 1951; *Trudy Fiziologicheskikh Laboratorii Akademika I. P. Pavlova*, 12 vols., 1924-1945; Gregory Razran, Extinction re-examined and re-analyzed: A new theory, *Psychol. Rev.*, in press.

[2] B. F. Skinner, *Science and Human Behavior*, 1953, 70; W. O. Jenkins, H. McFann, and F. L. Clayton, A methodological study of extinction following a periodic and continuous reinforcement, *J. Comp. & Physiol. Psychol.*, 43, 50, 155-167.

[3] Pavlov, [*Collected Works*], vol. 4, 401 f.

[4] Skinner, *op. cit.*, 102 f.

[5] J. S. Beritoff, [*Individually Acquired Activity of the Central Nervous System*], 1932; K. I. Platonov, [*The Formation of Motor Association-Reflexes to Simultaneous Auditory and Visual Stimuli*], Thesis, St. Petersburg, 1912.

[6] *Trudy Instituta Fiziologii Akademika J. Beritashvilli*, 5 vols., 1934-1943.

TWENTY-SEVENTH ANNUAL MEETING OF THE MIDWESTERN PSYCHOLOGICAL ASSOCIATION

The Midwestern Psychological Association met at the Hotel Sherman, Chicago, on April 28, 29, and 30, 1955. Northwestern University and the University of Chicago were host institutions. George S. Speer, of the Illinois Institute for Technology, served as chairman for local arrangements. A total of 1363 persons registered during the meeting. The program consisted of 131 papers and 5 symposia, emphasis being on learning, test construction and validation, and experimental studies in personality and social psychology. The symposia dealt with mutuality and reorientation, school psychology, advertising research, effects of the small group on personality, and the Iowa Picture Interpretation Test. The presidential address, "Wanted—a Good Cookbook," was delivered by Paul E. Meehl of the University of Minnesota. Dr. Meehl presented data and arguments to defend the position that actuarial methods of interpreting psychodiagnostic data were superior to clinical intuitive syntheses.

Benton J. Underwood of Northwestern University was elected president for 1955-1956, and Donald W. Fiske of the University of Chicago was elected for a three-year term as Secretary-Treasurer. New members of the Council are W. K. Estes, Indiana University, and George S. Speer. There were 284 new members elected, bringing the current membership to 1666. The 1956 meeting will be held on May 3, 4, and 5 at the Jefferson Hotel, St. Louis, under the sponsorship of Washington University.

University of Illinois LEE J. CRONBACH

FOURTEENTH ANNUAL MEETING OF THE CANADIAN PSYCHOLOGICAL ASSOCIATION

The fourteenth annual meeting of the Canadian Psychological Association was held in Halifax at Dalhousie University from June 2 to 4, 1955. Total registration was roughly 100. The program consisted of 16 papers on a variety of subjects, a symposium on training problems in clinical psychology, and a film on the sexual behavior of the laboratory monkey. W. Blatz and C. C. Pratt gave special addresses that were open to the public. Blatz talked on "Security in the child," and Pratt on "The meaning of music." The Canadian Navy arranged a cruise of the Halifax harbor for the occasion. The presidential address, given by Noel Mailloux, was concerned with the contributions of clinical psychology to the study of human behavior.

The following officers were elected for the ensuing year: President, George A. Ferguson, McGill University; President-Elect, Julian M. Blackburn, Queens University.

McGill University DALBIR BINDRA

FIRST ANNUAL MEETING OF THE SOUTHEASTERN PSYCHOLOGICAL ASSOCIATION

The Southeastern Psychological Association, which held its founding meeting in September 1954 in conjunction with the 62nd annual meeting of the American Psychological Association, met in Atlanta, Georgia, May 22-24, 1955, for its first annual meeting. Registration for the meeting totaled 268. The program listed 45 papers and four symposia. The first permanent officers are: President, John B. Wolfe, University of Mississippi; President-Elect, Nicholas Hobbs, George Peabody College for Teachers; Secretary-Treasurer, M. C. Langhorne, Emory University. Members-at-large on the Executive Committee are: C. H. Calhoon, Mental Health Consultant, USPHS, Atlanta; Arthur W. Combs, University of Florida; J. F. Dashiell, University of North Carolina. The Southeastern Association has a membership of 585 psychologists. The second annual meeting will be held in early May 1956 with Atlanta as its site.

Emory University M. C. LANGHORNE

George Howard Parker: 1864-1955

George Howard Parker, Professor of Zoölogy Emeritus at Harvard University, Fellow of the American Academy of Arts and Sciences, Member of the American Philosophical Society, and Member of the National Academy of Sciences, died on March 26, 1955, in Cambridge, Massachusetts, in his 91st year.

His contribution to psychology was considerable. The over-all problem with which he was concerned all his life was that of the origin of the nervous system. This led him to extensive comparative studies of sense-organs. His early research dealt with the compound eye of crustaceans. A little later he studied hearing and the functions of the lateral line organs in fish, and, later still, taste and smell in fish. In 1922 he published *Smell,*

Taste, and Allied Senses in the Vertebrates, a book which has been the standard source on that subject ever since.

Although the orientation of his research was always such as to interest the psychologist—the orientation had been worked out originally in the course of long discussions with William James—an estimate of Parker from the point of view of psychology alone—or indeed of biology alone—would fail to do him justice. He was an outstanding scientist not only in terms of research and of publications, but in terms of his teaching, of his relations with other scientists, here and abroad, and even in terms of the daily pattern of his life.

Born in Philadelphia December 23, 1864, he had by the age of ten years shown considerable interest in science, particularly in natural history. Explosions from youthful chemical experiments periodically upset the family household, as did the effluvia from fascinating, but rapidly decomposing, biological specimens. Attendance at Friends' Central School seemed to provide a good opportunity for the development of a budding scientist. Then, in the panic of 1877, financial calamity struck the family. Young Parker was forced to leave school to earn a living.

Instead of allowing this circumstance to terminate his education, he obtained the position of Jessup Fellow at the Philadelphia Academy of Natural Sciences. There, in the sort of environment which he loved, he managed to do not only the work for which he was being paid, but also to keep up with school subjects proper to his age. In consequence, he had no real difficulty in entering Harvard in 1883. Characteristically he ventured to Cambridge on $200, borrowed from a friend of his father, with no further financial prospects. Characteristically likewise, he rapidly repaid the loan and was thereafter wholly self-supporting.

Except for occasional special undertakings, Parker remained at Harvard for the rest of his life. He was Assistant in Zoölogy there 1886-1888, Instructor 1888-1899 (except for a trip abroad 1891-1893 as Parker Fellow), Assistant Professor 1899-1906, Professor 1906-1935, Professor Emeritus after 1935. Harvard was for him, however, a center of operations rather than a locus of operations. His trip abroad as Parker Fellow took him to many of the important zoölogical laboratories in Europe. Even before this, in 1888, he had begun what was to constitute a long series of visits to Woods Hole (he was Trustee of the Marine Biological Laboratory there 1907-1935, Emeritus Trustee thereafter). After election to the National Academy in 1911, he undertook for the Federal Government a mission to the Pribiloff Islands in the Bering Sea (to "count seals," as he himself put

it). In 1926 he was delegate from Harvard to the Pan-Pacific Science Congress in Tokyo, and visited not only a number of Japanese universities but also several in China.

Both at home and on his travels Parker had the gift of establishing close and lasting relations with fellow-scientists. His delightful autobiography, *The World Expands* (1946), is largely a journal of meetings and doings with the notable biologists of the period from 1890 to 1940. Alexander Agassiz, Asa Gray, William Morton Wheeler, E. G. Conklin, E. B. Holt, W. J. Osterhout, W. B. Cannon, A. V. Hill, Hans Spemann, E. D. Adrian, A. Szent Gyorgi all were his masters, friends, or students. His soundness as a scientist, breadth of interests, clarity of exposition, considerateness as a host or guest (with the constant help of his wife, the former Louise Merritt Stabler, whom he married in 1894), even readiness to join in larks (drinking home-made red wine out of coffee cups in his favorite Italian restaurant during prohibition)—all these were surely involved. Whatever the causes, however, the result was that in an age where science is the joint undertaking of large and highly organized groups his influence in the world of science was tremendous for near on half a century.

With the death of George Howard Parker, psychology has lost an important contributor, zoölogy an outstanding authority on the development of the nervous system, and the scientific profession one of its most eminent, most influential, and most beloved members.

Harvard University J. G. BEEBE-CENTER

Godfrey Hilton Thomson: 1881-1955

Godfrey Thomson's death on February 9, 1955, represents a great loss to psychology, though his early training and research were in mathematics and physics and his only strictly psychological appointment was a year spent with E. L. Thorndike at Columbia. Graduating from what is now King's College in Newcastle-upon-Tyne, he took his first doctorate *summa cum laude* at Strasbourg in physical science. The conditions attached to a fellowship he had held from his old university led him to accept an appointment there as lecturer in education. He became Professor in 1920, and went to Edinburgh as Professor of Education and Director of Moray House Teacher's Training College in 1925.

Out of his work in education grew an interest in psychology which was supplemented by a long vacation spent with C. S. Myers in the Cambridge laboratory. At first, as one would expect from a physical scientist, labora-

tory methods occupied his attention, and he published several papers in experimental psychology dealing particularly with psychophysics. With the appearence of Spearman's two-factor theory before the first World War, a final change in Thomson's orientation took place. His empirical and critical mind reacted against Spearman's rather sweeping claims, and he set about to show how they must be qualified. Two important papers resulted, one in the *British Journal of Psychology* for 1916, and the other in the *Proceedings of the Royal Society* for 1919. There ensued one of the most spirited and at times heated controversies in the history of British psychology. It still makes good reading, and reveals much of the personalities of both men: Spearman—erudite, speculative, and dogmatic; Thomson—ingenious, sceptical, and pugnacious. Theoretically, they were never reconciled, though each came to hold the other in high esteem.

The gist of Thomson's argument was that a correlation matrix of low rank can be derived from the chance interaction of an infinitely large number of small influences, and that Spearman's two-factor solution, while possible, is not necessary. The same skeptical attitude towards factorial methods Thomson later extended to L. L. Thurstone's multiple-factor analysis, but it was a tempered skepticism, admitting that Thurstone might be right, but insisting always that other possibilities must be kept in mind. Thomson's book, *The Factorial Analysis of Human Abilities* (1939), is one of the classical works in the field.

While the theory of factor analysis occupied first interest for Thomson, he contributed much to other areas of mental measurement. A practical problem after his own heart was set him by Northumberland Education Authority when they asked him to devise the fairest possible method for selecting children who were to have free secondary education. This demand led to the Northumberland Tests, now the Moray House Tests, and to what Thomson—who would himself have been compelled to leave school at thirteen had he not won a series of scholarships—described as "a lifelong task I felt bound to persevere in for the sake of intelligent children." Nowadays the Moray House Tests are used each year throughout Britain and have served as models in many parts of the world. In Britain alone the academic careers of more than a million children are affected each year by these tests.

Analysis of results suggested that there might be a negative correlation between test-score and size of family, a finding which induced much speculation concerning a possible lowering of the intelligence of the population. A series of studies to check on this point were initiated in

1946, and preliminary findings were reported in 1949. It was hoped to follow up some of the specially tested children for twenty years. As was typical of Thomson, the plans for these studies were scientifically thorough. Instead of being satisfied with easily available groups of children, he insisted upon a representative sample containing every Scottish child born on the first day of each month during 1936. These studies are still continuing as planned.

So that he might, as he put it, "safeguard himself from the temptation to make money out of this activity," Thomson established a trust to administer the very considerable income brought in by the Moray House Tests. The money has been used to further educational research conducted by a strong team of investigators at Moray House.

Thomson was also a brilliant practitioner of the educator's art. Although it is difficult to find a yardstick by which the success of a teacher can be measured, we may note that his students occupy chairs of psychology or education in half a dozen British universities and as many again have similar positions overseas. Many psychologists now in key administrative positions were also trained by him.

Honors have not been lacking. In addition to the knighthood conferred upon him in 1949 and a number of British academic distinctions, Thomson was an Honorary Member of the American Academy of Arts and Sciences, a Foreign Associate of the National Academy of Science, a Fellow of the American Association for the Advancement of Science, and a Member of the Institut International de Statistique.

Thomson will be remembered for his contributions, both theoretical and practical, to psychology. On the theoretical side he reacted against the sweeping claims initially made for the two-factor method by presenting mathematical developments showing that other interpretations were possible and deserved careful consideration. He combined his mathematical skill with a strong belief in the necessity for psychological measurement and with a common-sense factual approach to problems. This approach led to his major positive contributions and brought about a wide scale application of testing methods in the educational field. The Moray House Tests which he initiated and firmly established will continue to assist in the proper academic placement of millions of children.

University of Edinburgh JAMES DREVER

BOOK REVIEWS

Edited by M. E. BITTERMAN, The Institute for Advanced Study, Princeton, New Jersey

Experimental Psychology. By ROBERT S. WOODWORTH and HAROLD SCHLOSBERG. Revised Edition. New York, Henry Holt and Company, 1954. Pp. xi, 948.

In spite of the appearance of a handbook and several textbooks of experimental psychology in the past few years, this revision of Woodworth's 1938 text has been eagerly awaited. The readability of the earlier book, and Woodworth's happy faculty of presenting the facts in a background that gave relatedness and meaning to the whole, were doubtless responsible for its long domination of the field. His respect for the work of psychologists of other nationalities and for work yielding qualitative as well as quantitative results were other virtues. The question is whether the distinction of the earlier book could be maintained in the face of the ever-increasing tide of experimental literature of the last 15 years. Could the narrative quality and the impartial presentation of results, qualitative or quantitative, American or European, be preserved without omissions which some reader would surely question? Certainly no current writers are better prepared to tackle this task than are Woodworth and his new collaborator, Schlosberg. They have tried very hard to keep the virtues of the old book, with all the success that could be expected today. The preface points out that over 50% of the articles and books now cited (2480) were not used in the old edition. A high proportion of these appeared after 1940. Sacrifices of the old material have had to be made. The qualitative data seem to be less well represented; the chapter on memory for form, one of the real highlights of the old book, is gone and only partially covered in another chapter. These are trends of the times, and new chapters, such as the one on discriminative learning, have superseded the old.

The style of the book reflects a choice which had to be made between a textbook and a handbook type of presentation. The new volume is still a textbook, easy, direct, and addressed to the reader. Most of the chapters get off to a gentle start, sometimes with homespun examples and phraseology (in the first chapter on learning, the distinction between a "problem" and a "lesson"). Sometimes a section ends with a summary statement which is a little too easy a generalization, *e.g.* "The learner never starts from scratch but always with a handicap" (p. 604), or "The animal learns what stimuli are coming through classical conditioning, and then finds out what to do about them by instrumental conditioning" (p. 554). The core of each chapter is, however, always packed with information, often at a high level of difficulty.

As the authors point out in the preface, they have espoused no systematic viewpoint, but have tried to maintain an eclectic approach. The seriousness of this intent is supported by a later statement: "the reader must realize by this time that topics and chapters are no more than conventional groupings of related aspects of behavior—merely arbitrary categories in a filing system" (pp. 151 f.). Nevertheless, the functional approach, as the authors suggest, is still apparent. This approach is aptly illustrated in their discussion of a stimulus, which is defined as a physical or

chemical agent, but "may also be regarded as a signal. It is a signal, psychologically, because the motor response is adjusted to the environment that provided the stimulus" (p. 267). Further evidence of the functional approach is the formula familiar to all users of Woodworth's introductory textbooks, $R = f(S, O)$, which is used throughout the book as a classificatory and descriptive aid. In the chapter on attention, for instance, the types of experiment are classified by variations of this formula. Experiments on fluctuation of attention are formulated as $R = f(O, t, S_1, S_2, S_3, \ldots)$, divided attention as $R_1R_2 = f(O_1O_2, S_1S_2)$, and so on. In the reviewer's opinion, the space occupied by these schematic descriptions might better have been used for presentation of more experiments or for discussion of theories. Definitions, on the whole, are operational. This term is explained somewhere, but it would be convenient if it were explained in the first chapter, as well as the terms "hypothetical construct" and "intervening variable" which are used without explanation.

The book begins with a short introductory chapter which is concerned with the variables of an experiment. Experimental design is not treated; instead, an extensive section on methods precedes directly each of the two principal sections of the book, the eight chapters devoted to sensory and perceptual processes and the nine chapters devoted to learning and problem solving. Each chapter contains, as well, technical material relevant to the investigation of its topic, and the methods are made clear and practical for laboratory use.

Following the introduction, there are three chapters devoted to topics which held priority in the history of psychology—reaction-time, association, and attention. The chapter on reaction-time gives considerable space to techniques of measurement. Here is introduced an explanation of logarithmic scales and a useful discussion of fitting an equation to data. The chapter on association contains, in a discussion of distributions of association reaction-times, a handsome comparison of four methods of plotting, illustrating graphic tests of normality. The chapter on attention brings the description of tachistoscopes up to date and has a very complete review of experiments done with them, from Dodge to "subitizing" and "perceptual defense." The authors suggest with regard to attention that the psychologist avoid subjectivity and keep his bearings by taking as his model a cat waiting at a mouse-hole. Would this admonition really be so relevant for understanding tachistoscopic viewing and estimation of numerosity?

The three chapters on emotion which follow begin with a review of theories, one of which, an "activation level" theory, is chosen as a means of organizing the experimental material. The bodily changes characteristic of emotional states are treated as indices of activation level; the *GSR*, though not treated in a separate chapter, comes into its own with 23 pages of results and method. The most interesting part of these chapters (to the reviewer) is Schlosberg's own effort to scale emotional expression. It serves not only to present information, but as a nice methodological example of working out a scale. The most unsatisfactory part, for some psychologists, will be the closing statement that all the work on bodily changes has been worth while because it gave us the lie detector. "Lie detecting puts level of activation to work!" (p. 191).

The two chapters on psychophysics are considerably changed from the old book, which devoted one to methods and one to results. They are divided now into "Determination of Thresholds" and "Scaling Methods." The reviewer always found

the two chapters in the old book extremely useful, and they are now equally useful. The methods are clearly presented with plenty of illustrations and examples. The use of z-scores for determining threshold values with a constant method is explained very thoroughly, and arguments for and against Müller-Urban weights, three categories, and like problems are presented in such a way as to be helpful rather than confusing to a student trying to decide which procedure to follow. Blackwell's work on forced choice is not presented. The single-stimulus method is happily allotted a fair share of space. The discussion of constant errors and series effects is very interesting, though there is a sticky place where the authors take up compression of the frame of reference. The use of the symbol *Co* for comparison-stimulus is a little bothersome since it is so easy to read it as *constant*. The trace concepts used in this section (fading, sinking, assimilating) are rather non-operational in view of the tenor of most of the book. One wonders if it would be possible to keep traces out of psychophysics and treat judgment consistently as behavior.

The chapter on scaling describes both psychophysical and psychometric scales and points out the similarities of method. This chapter gives the reader a feeling that psychology has really got somewhere. Something must come out of these procedures which is scientifically worthy. What does is presented in the next six chapters on sensory processes.

These chapters read more like a handbook than the other sections, but the point of the experiments described is always clear and any student who takes a little trouble will find himself deeply interested. A good deal of technical detail about control of stimulation is given for the cutaneous senses and especially for smell; the nature of the stimulus for smell is also discussed at a technical level. This treatment should serve well the very important point of teaching the student once and for all that stimuli and objects are not the same thing. The receptor structures and processes are treated in some detail, but the neural pathways are barely mentioned. Of course the chapters are already very full, but audition and vision particularly seem to be cases where a little more information on neural pathways and the centers involved would have been interesting, especially since a number of recently studied phenomena, such as binaural effects, are referred to the centers. The chapter on vision contains an impressive proportion of new material, such as information on the human electroretinogram, the quantum theory of the visual threshold, and evidence for a four-component theory of color vision. In this chapter, as in audition, the authors point out the "current trend to push analysis and synthesis back from the receptor to the higher centers" (p. 402). In spite of this comment, which implies that central processes are critical even for sensation, the chapters on perception begin with a distinction between sensation and perception. For perception, the student is told, "something more than reception of the stimuli must occur" (p. 403).

It is notable, in considering the changes which have occurred in the past 15 years, that the old chapter on perception of form introduced the term Gestalt in the first sentence. In the new chapter with the same title, the word never appears. This is not to say that Gestalt theory is not given its due—Wertheimer's laws are still quoted and Köhler's satiation theory gets several pages. In the latter connection, it would have been good to find a report of Ivo Kohler's extensive experiments on prolonged adaptation to, and after-effects of, optical distortion. The chapter on perception of color is not much changed, except for a few additions such as an

intriguing bit on the relation of color-constancy and space. The chapter on visual depth-perception is entirely rewritten. It includes an excellent discussion of stereoscopes, a long section on size-constancy (with a comparison of various kinds of size), and a description of a number of Ames' demonstrations. The last of this group of chapters is a fine one on eye-movements in relation to perception. The types of eye-movement are described, as in the old book, but now they are integrated with reading, the perception of motion, and postural and visual orientation in space. There seems to be a dearth of experiments on the perception of real movement—none are mentioned, though the problem is raised. In all these chapters, nothing is said about the perception of time.

The remainder of the book (nine chapters) is devoted to learning. The first of these not only discusses methods, but makes clear the middle-of-the-road position which will be taken: "habit strength" and "memory trace" are declared operationally equivalent. (Can hypothetical constructs be operationally equivalent?) The chapter on conditioning is entirely different from that in the old book. The "law of reinforcement" and the "law of extinction" are stated *à la* Skinner. The statement is said to be operational, but since the terms "unconditional reflex" and "conditional reflex" are used in the statement and are not defined (except in terms of strength) not much may be gained. In any event, the rest of the chapter is theoretically oriented and there is a comprehensive discussion of the nature of reinforcement and of extinction. There is a good-sized section on sensitization and pseudoconditioning; sensory pre-conditioning is omitted. The section on generalization and the form of the typical gradient (if any) is admirable.

The new chapter on discriminative learning, like the rest of the book, avoids definition and states the problem by example (a rat in a discrimination-box). What the rat has to learn in this situation is discussed. He "has to acquire an 'association' between the bright stimulus and the positive response" (p. 583). There is no further discussion of the nature of the response in the discrimination-box, but the parallel problem is given 20 pages of discussion in the following chapter on maze-learning (does the animal learn turns or does he learn the environment?). The typical problems of discriminative learning, equivalence of stimuli, continuity and non-continuity theories, are well reported, though statistical theories of discriminative learning are not mentioned, nor is Gulliksen. At the end of the chapter there is a section on human concept-formation. There are surely some good reasons for putting it here—most recent work on concept-formation tries to relate it to discrimination—but the reasons are not clear. Somebody should formulate them soon.

"Motivation in Learning and Performance" is a new chapter and a welcome addition. The physiological basis of needs and drives is not so much emphasized as are the more psychological problems, such as secondary motives, punishment, frustration, motivation of work, likes and dislikes. There is a section on generalization of drives and incentives which faces us with the difficulties of using this concept too liberally. How are we going to define generalization here? As equivalence of drives or incentives? But the gradient has always, presumably, been a principal characteristic of generalization. As the opposite of generalization? Differentiation is learned and generalization is not, the authors say. But then there are equivalent secondary incentives, so learning must be involved. It seems clear that this is a confused concept which needs new and strict defining. The authors point out, laudably, the

"easy-going plausibility of such explanations," and that "drive generalization and incentive generalization are a pair of very handy conceptual tools for any psychologist who wishes to derive all human motivation from a few basic organic needs" (p. 683).

The chapter on memory begins like the old one, but omits some of the classical experiments. The reader is referred to the previous edition. Learning of visual forms is included here, as is the topic of retention. Theories of retention, however, are saved for the next chapter on transfer and interference, where they follow theories of retroactive inhibition. This seems a very effective sequence. Woodworth's useful distinction between transfer and transfer-effect is still in this chapter. The section on transfer in terms of *S* and *R* is much extended and now includes some paragraphs on transfer of cue-differentiation which may set experimenters working on this important problem. Economy of learning and performance come next and introduce a new section on work and fatigue, which is brief but critical and well-implemented. The final chapter on problem solving contains much new material—less on animals, more on human *S*s. The old and profitless trial-and-error-vs.-insight controversy is no longer the theme of the chapter, which gives more emphasis to the rôle of variables such as previous experience, information, and set. The conclusion points out the promise of this field and its hopeful future.

Among the noteworthy physical features of the book are the excellent illustrations. They are numerous, mostly graphic or diagrammatic. They always add to the text and are supplied with longish explanations. The end-papers are printed with log- and probability-scales and with a table of 4-place logarithms. Whether they prove useful or not, they look attractive. The bibliography serves as an author-index and now contains titles.

In the final analysis, this reviewer is awed by the magnitude of the job that Woodworth and Schlosberg have done. All of the critical comments here made are mere plucking at straws. The book will surely rival the old one in winning the loyalty of generations of students and the gratitude of their instructors.

ELEANOR J. GIBSON

Cornell University

Personality Through Perception. By H. A. WITKIN, H. B. LEWIS, M. HERTZMAN, K. MACHOVER, P. BRETNALL MEISSNER, and S. WAPNER. New York, Harper & Brothers, 1954. Pp. xxvi, 571.

The study reported here is part of a comprehensive research program that has been in progress under Witkin's direction since 1942. The very first problem was to determine the factors responsible for the maintenance of proper orientation to the upright in space and quickly led to a set of ingenious experiments reported in 1948 by Witkin and Asch. From these investigations it was firmly established that, when a 'strong' visual field is present, the perceived upright is determined with relation both to the axes of that field and to impressions received from the body, with visual factors playing the dominant rôle. Most striking were the wide and highly consistent individual differences observed in the extent to which *S* depended upon visual rather than kinaesthetic functions when the two sources of information were in conflict. At one extreme were individuals who relied almost exclusively on the visual field. At the other were *S*s who relied almost entirely upon bodily experiences, disregarding the visual field.

It was because of this great variation from one person to the next that Witkin turned his attention from general laws to the problem of determining why such marked individual differences in field-dependency exist. The series of experiments herein reported deal with three major problems: (*1*) the pervasiveness of field-dependency among different tests of spatial orientation, its bearing upon other tests of perception and motor skills, and its stability through time; (*2*) the developmental sequence of factors entering into spatial orientation from childhood to adulthood; and (*3*) the relation between an individual's characteristic way of perceiving and his general personality organization.

In the major study of 52 men and 51 women, who were psychology students at Brooklyn College, measures of ability to perceive the true vertical were obtained from three variations of the rod-and-frame test, five series of the tilting-room-tilting-chair test, and three subtests using a rotating room. In each of these tests the *S* could locate the upright according to the axes of the visual field or with reference to bodily sensations, and he indicated this location by adjusting his body, the field, or the rod to a position he perceived as vertical. Related perceptual and "body-action" tests consisted of an embedded-figures test of the paper-and-pencil type adapted from Gottschaldt; a test of brightness-constancy; a test involving auditory-visual conflict; a body-steadiness test (ataxiameter); a body-balance test (stabilometer); and a two-hand coöordination test (a standard pursuitmeter). Personality measures were obtained from an autobiography, the *MMPI*, a sentence-completion test, a two-hour interview, the Figure-Drawing Test, the Rorschach, the *TAT*, and a word-association test.

Besides the college students, two other groups were studied: 38 male and 39 female patients from the psychiatric wards of Kings County Hospital, and 270 children in five age-groups ranging from 8 to 17 years. Each psychiatric patient was given two of the orientation tests, the embedded-figures test, and the same personality tests used for the college students. The three perceptual tests were also used in the developmental study employing children, thus making it possible to compare three different populations with respect to degree of dependence upon the visual field. Personality measures for the children were obtained from the Rorschach, the *TAT*, the Figure-Drawing Test, and a miniature-toy play-situation.

The generality of field-dependence among the different orientation tests and the other perceptual-motor tasks was determined by analysis of intercorrelations, first for the subtests within each of the three orientation situations, then for composite scores among the three orientation tests, and finally for a general orientation-index and the perceptual-motor tests. Some appreciable correlations were found between measures of field-dependency in vertical orientation and each of the other tests, with the exception of the task involving auditory-visual conflict, thus substantiating the hypothesis that field-dependency is a fairly general trait. Particularly striking is the correlation between the over-all orientation-index and scores on the embedded-figures test—.66 for men and .46 for women. Witkin cites a factor analysis, done independently by Linton, in which the embedded-figures test was the most highly saturated of any test on the major factor (ability to keep an item separate from field). If this relationship continues to hold up, the embedded-figures test should prove to be a highly valuable measure of field-dependency since it requires no special apparatus and can be given quickly.

Some interesting differences between the sexes appeared, especially in adults. Generally speaking, women proved to be much more dependent upon the visual field and less consistent from one test to the next than were the men. Analysis of perceptual functioning in children of different ages revealed the same sex-difference, though to a lesser extent. When the orientation subtests were divided into those which featured perception of an item within a field and those which involved perception of the field as a whole, different developmental curves were obtained, especially between the ages of 10 and 13.

Witkin's major objective throughout the rest of the book was to demonstrate relationships between perceptual functioning and a number of personality characteristics. To accomplish this, he teamed up with several clinical psychologists who worked semi-independently of him in the gathering and analysis of the data. Hertzman was responsible for the Rorschach and *TAT;* Machover handled the Figure-Drawing Test; Meissner worked with the play-test; Lewis conducted the interviews; and Lewis and Witkin shared responsibility for over-all integration of the perceptual and personality data. Wapner worked primarily on the self-consistency studies and the statistical analyses. In a coöperative project involving so many people, uneveness of style and disjointed sections are hard to eliminate, and the present book is no exception.

Correlational techniques were used as the main procedure for comparing perception and personality. In addition, the *Ss* were divided into three groups, high, medium, and low, according to their orientation scores, and detailed analyses were made of group trends. Almost invariably, the latter method resulted in redundancy and over-analysis of the data, often to the point where minor trends which could best be interpreted as chance fluctuations in small samples were declared significant for purposes of proving an hypothesis.

The most striking correlations were reported for clinical ratings based upon combined data from the interview, the autobiography, the sentence-completion test, and a personality questionnaire. The ratings were made by the interviewer in five areas: self-awareness, hostility, activity-passivity, inferiority feelings, and self-assurance. Some of the individual correlations between the personality ratings and the orientation scores were as high as .70, a rather startling finding when the probable unreliability of single clinical ratings (often below .60) is taken into account.

Correlations between the Rorschach scores and the perceptual data are likewise high, and in this case there is good reason to suspect that they are spuriously high. In selecting variables and cutting-points for scoring the Rorschach, six protocols of high-index performers on the orientation tests and six low-index performers were carefully studied to eliminate all but the most promising variables. To the extent that this pre-comparison with the criterion exploited chance factors, the resulting correlations are spurious. Some cross-validational data are presented for the hospital group in which only two of the variables, those having to do with location and popularity of the Rorschach responses (*W* and *P*), still correlate significantly with the orientation-index; but since this population is entirely different, it doesn't provide an adequate test of the validity of the reported relationships in normal adults.

Witkin and his colleagues repeatedly fall into the grave error of accepting pat 'clinical' interpretations of test-data as valid in spite of the fact that few, if any, have proven correct in previous experimental work. For example, the *M, H, F,* and

A scores on the Rorschach are declared to "directly reflect the inner life of the person, being related to self-awareness, fantasy, and to some extent self-acceptance" (p. 206); while *W*, *P*, and *C* are lumped together as measures of "organizing ability, activity level and capacity for control of impulses" (p. 206). Machover's interpretations of the Figure-Drawing Test are even more highly speculative, though presented with an air of authority. Such claims to validity for the personality data would perhaps be less objectionable if there were any evidence to indicate high agreement among clinicians working independently, but nowhere is there concern about the probable unreliability of the personality variables employed.

In an effort to prove the general hypothesis that there is a strong relationship between field-dependency as defined by the tests of spatial orientation and a variety of personality measures, the authors grow more and more reckless in making unwarranted inferences as they present additional data from the college and hospital samples. One entire chapter is comprised of six select case "histories" drawn from the college sample. For the most part, these presentations are intuitive, highly speculative, psychoanalytically-oriented interpretations with little or no case-history data of a factual sort. Yet the authors blandly assert that "A comparison of the personality patterns of these six subjects makes clear certain aspects of personality differences among people with different modes of perception" (p. 312), and that "The differences in personality . . . confirm the thesis that perceptual performance and personality are related" (p. 314).

After examining case records for high- and low-index performers in the sample of male patients, the authors conclude that "Open castration wishes, weak impulse control, crushed enfeebled self-attitudes, tend to go along with field-dependent perceptual performance." Examination of pertinent data summarized in their Table 15.6 reveals that such a conclusion is wholly unwarranted. For example, "open castration wishes" are part of the symptom-picture for only three of the 14 high-index men, two of whom were ascribed this wish because they desired to be circumcized; furthermore, "weak impulse control" and "crushed enfeebled self-attitudes" are actually present in the symptom-pictures of low-index males about as often as in the field-dependent men. Such careless disregard for reasonable caution in the interpretation of data in many sections of the book does little to build up the reader's confidence in the ability of the authors to analyze and report data in an objective, scientific manner.

In spite of numerous speculative inferences and methodological weaknesses, many of the findings concerning the relationships of field-dependency to personality cannot be lightly dismissed. The unusually high correlations between interview-ratings and perceptual scores, the repeatedly significant relationships between field-dependency and certain of the more objective Rorschach scores, and the apparent ability of the clinician to rank "blindly" 25 randomly selected figure-drawings as to degree of field-dependency in the perceptual tests (only 4 *S*s misranked), are only a few of the findings worthy of serious consideration. If they have not satisfactorily demonstrated the general validity of their major thesis, the authors have at least presented an extensive array of stimulating hypotheses for further testing by other investigators.

University of Texas

WAYNE H. HOLTZMAN

Decision Processes. Edited by R. M. THRALL, C. H. COOMBS, and R. L. DAVIS. New York, John Wiley and Sons, 1954. Pp. viii, 332.

In 1944, von Neumann and Morgenstern published their well known book on the theory of games. Since then, stimulated primarily by that book and secondarily by the work of Wald and others, a new interdisciplinary study of decision making has grown up, in which mathematicians, statisticians, economists, psychologists, and others are participating. The problems with which they are concerned might be put as follows: if a person, group, or institution is faced with two or more alternative courses of action, and if it knows something (but not necessarily everything) about the various possible outcomes of each course of action and the way in which these outcomes are contingent on various possible states of the environment, then (*a*) which course of action should be chosen? and (*b*) which course of action is in fact chosen in real-life situations.

The reader probably will recognize (*a*) as an ethical problem and (*b*) as a psychological problem. The new field is characterized, not only by a concern with these problems, but also by a method of attack. That method is to set up one or more matrices, each cell of which represents the payoff accruing to the decider as the result of a particular decision in a particular state of the world, and then to attempt to answer questions by means of mathematical operations on various kinds of metric and semi-metric information entered therein. All of the chapters in *Decision Processes* are concerned with this method of attack, explicitly or implicitly. Since it makes choice of a course of action hinge on the returns of that course of action, the method is based on a kind of hedonism, which is nothing against it, of course, as long as it leads to correct predictions.

In 1952, under primary sponsorship of the Ford Foundation and the University of Michigan, an eight-week conference on problems and experiments in decision-processes was held at Santa Monica. This book grows out of that conference. Some of its papers were presented at the conference, some grew out of work started at it, and some are by people working in the same area who did not attend but might well have. Let me say at once that this is a top-notch book. It is easy to call it the best in its field since it is the only one, but even if it had dozens of competitors, it might well be the best. Multi-authored books of this consistently high quality are rare.

The first chapter, by Davis, is an excellent review and summary of the whole book. The second chapter, by Coombs, Raiffa, and Thrall, summarizes the relation between mathematical models and data, using measurement-theory as example. (This chapter has been published elsewhere in psychological literature, as has Chapter 16, by Hoffman, Festinger, and Lawrence, and the substance of Chapter 9, by Estes. The rest, so far as I know, are new). Of the remaining chapters, six are concerned with a single mathematical model for decision-making which supposes that *Ss* choose in uncertain situations in such a fashion as to maximize the product of two mathematical entities called *utility* and *subjective probability*. Four chapters are concerned with the experimental and theoretical development of statistical learning theories of the Bush-Mosteller and Estes type. Three chapters are concerned with the question of developing rules for decision-making which reflect in some way the tastes and wishes of the members of a group (simple majority vote is only one of many possible rules). Two particularly illuminating chapters discuss various possible rules for

decision-making in general, showing the relationships, agreements, and differences among these rules. The remaining two chapters present experiments on game-playing which are directly related to certain issues in the theory of games.

It would take far too much space to summarize the chapters in detail, but the introduction to the book does that very well. I will only mention here certain chapters which seem to me to be especially valuable. Chapter 4, by Milnor, summarizes four decision-criteria which have been suggested in the literature, shows that they are inconsistent with one another, examines the properties for which they were selected, chooses a particular set of properties which seem most desirable, and constructs a new (but hard-to-use) criterion having these properties. I found it one of the two most illuminating chapters in the book. The other, Chapter 17, by Coombs and Beardslee, presents a most ingenious graphic analysis of the utility-subjective probability model, an analysis which must be clear to anyone regardless of mathematical sophistication, and which highlights the features of the model as no other publication in the literature has done. Mathematicians tell me that Chapter 8, by Bush, Mosteller, and Thompson, which gives a very abstract treatment of the Bush-Mosteller learning model plus some new developments dependent on a restrictive assumption, and Chapter 12, by Hausner, which presents a mathematical theory of multidimensional utility-functions, are especially exciting. I cannot confirm or deny this opinion; my mathematical skills are not great enough. Economists are excited by the work on teams, foundations, and coalitions which Marschak presents in Chapter 14.

The dominant notion of this book, and of the field of decision-making, is that of the utility, or subjective value, of objects, events, and amounts of money. The greatest weakness of this book taken as a whole is that it gives no adequate answer to the question of how to go about measuring utility. The only method of measurement which is explicitly presented is based on a very restrictive and unlikely assumption and, even so, gives only an ordered metric (less than an interval scale). In other chapters, in which the notion of utility is used empirically, the assumption is made that utility is identical with monetary value. Everyone recognizes that this assumption is improbable in the absence of evidence, and the fact that it must be made is not the fault of the authors of the book. Although utility is, and is likely to remain, the central concept in the area of decision-making, all methods of utility measurement which have been attempted experimentally are singularly unsatisfying (including one not yet published by this reviewer). Eventually someone is going to have to do something about this problem.

The book reports the results of 15 experiments (a number which is a bit arbitrary, since it is hard to decide when one experiment ends and another begins). Of these, two are summaries by Estes of experiments which have been fully reported elsewhere, and one, to which the entire chapter by Hoffman, Festinger, and Lawrence is devoted, has been published in essentially identical form elsewhere. This book provides, as far as I know, the primary publication of the other 12 experiments. Of those 12, only one (by Coombs) meets reasonable minimum standards of design, execution, and reporting. Unpublished extended accounts of them lead me to believe that the others fail to meet minimum standards of design and execution. Of course, the authors all recognize this fact. They call their experiments 'pilot' experiments, and they are careful to point out their limitations. Nevertheless, they present data and

draw conclusions. The theory of games lends itself to the design of informal experiments which are conducted on graduate students, family members, secretaries, or colleagues. The Santa Monica Conference was a hotbed of such studies. I do not believe, however, that the results of such enterprises, invalidated as they are by the choice of *S*s and general sloppiness of design, should be published to waste the time and puzzle the brains of unwary readers. There is the most urgent need in the area of decision-making for good experiments to interact with the available wealth of good theoretical work. Lest this criticism scare you away from the book, I should add that relatively little space is taken up by the presentation of these 11 experiments.

The book would have profited from a considerably sterner exercise of the editors' blue pencils. If a concept appears in two different chapters, it generally is represented by two different symbols. Three of the chapters (7, 15, and 16) are unreasonably wordy for what they have to say, and one of the three (15) is so divergent in tone, content, and purpose from the rest of the book that it should probably not have been included. Four of the chapters (8, 10, 11, 14) are filled with mathematical symbols and abstract formulations beyond the demands of content. One chapter (19) presents some simple experiments on game-theory, but discusses them in terms which are unexplained and unknown except to game-theory experts. In general, the level of sophistication in mathematics, game-theory, and statistical decision-theory which is needed varies very widely from chapter to chapter. Five are understandable to mathematical illiterates. Nine more have about the same difficulty as an undergraduate psychological text in statistics. The remaining five are rather difficult—three of them in my opinion unnecessarily so. (The editors give their own judgment of mathematical difficulty in the Introduction; I do not completely agree.)

Wiley is to be congratulated for publishing this book, but not for the method of publication. The unjustified right margin looks untidy, the occasional hand-drawn symbols are unclear and confusing, and the unavailability of italics makes for confusion between symbols and text. The number of typographical errors is, however, gratifyingly low.

If I have not persuaded you by now to buy this book, I am unlikely to be able to do in a final paragraph. I would like to. It is a splendid book in a new field which looks as if it will become increasingly important in psychology. As such, it belongs on the shelves of any psychologist who is, as the Psychometric Society puts it, interested in "the development of psychology as a quantitative rational science."

WARD EDWARDS

Air Force Personnel and Training Research Center

The Six Schizophrenias: Reaction Patterns in Children and Adults. By SAMUEL J. BECK (with a clinical introduction by ROY R. GRINKER and a chapter by WILLIAM STEPHENSON). New York, The American Orthopsychiatric Association, 1954. Pp. viii, 238.

This monograph reports an investigation which valiantly attempts to provide a basis for breaking away from Kraepelinian nosology. Whether success is achieved is uncertain. One's evaluation depends upon one's faith in the psychologists' and psychiatrists' *Q*-sorts—the crude ore of the study—and in the special statistical methodology which refined it.

Briefly, the study was conducted as follows: A large pool of clinical items, presumably characteristic of schizophrenia, was assembled by the psychiatric staff and residents of Michael Reese Hospital. The items were grouped, somewhat arbitrarily, into four classes—defenses, ego-processes, emotional forces, and restitutional forces. For most of the clinical items, Rorschach correlates were specified. An example of a clinically-formulated item is: II a. "withdrawal from social contact, but not into autistic fantasy." Its Rorschach correlate is: "Animal percent low with unique and original interest content; without M. In some, preoccupation with one interest; as sex, or anatomy. Stereotypy, with narrowed range of interest, in others. In them, animal percent may be high" (p. 55).

The clinical outline was converted (by Stephenson) into a *Q*-sort deck of 120 items of the type illustrated by the following: "inadequate reality testing, distorts external stimuli," "has little energy," or "excessively submissive." For each of 20 schizophrenic patients (12 adults and 8 children) four *Q*-sorts were made, one by a psychiatrist on the basis of anamnesis, interview, and other psychiatric data, and three by psychologists on the basis of Rorschach protocols. Since two of the psychologists had high agreement with one another, the *Q*-sorts of one were dropped from the matrix. The statistical analysis, then, was completed on the basis of the intercorrelations of the *Q*-sorts of a psychiatrist (not always the same one) and (in most cases) two psychologists. Since the complete matrix would have been two cumbersome, two smaller matrices were constructed. The first represented the 12 cases where psychiatrists and psychologists agreed. From the 465 coefficients, a factor analysis yielded three schizophrenic reaction-patterns or types. The second matrix was made up of the remaining eight cases where the psychiatrist and the psychologists disagreed in their *Q*-sort assessments. From the 351 coefficients in this matrix, two additional reaction-types were derived. These were identified only through *Q*-sorts based upon Rorschach protocols. An extension of the study involving 20 children produced a sixth reaction type, a changing and apparently transient form of schizophrenia.

Further analysis revealed the factorial structure of the six schizophrenias. Five factors were isolated, each with two or three "levels of manifestation." The first factor is identified as *defense organization;* its three levels of manifestation may be assessed as "withdrawal," "constriction," and "pathogenic." The second factor is labeled *intellectual functioning* and may be "orderly" or "disrupted." The third factor is labeled *fantasy activity* and may be "autistic," "regressive," or "little or none." The fourth factor, *social adaptation,* has three levels, "self-absorption," "self-deprivation," and "restitutional." The fifth factor, *emotional state,* may be characterized by "lability" or "fixed tone."

Needless to say, the factors are not the types. Reaction-type *S-1,* for example, is characterized by a *pathogenic* defense-system, *disrupted* intellectual functioning, *little or no* fantasy activity, and *self-absorptive* and *restitutional* social adaptation. Reaction-type *SR-2* (identified by Rorschach) is characterized by *withdrawal* and *pathogenic* defense organization, *orderly* intellectual functioning, *autistic* and *little or no* fantasy, *self-absorptive* social adaptation, *labile* emotional state, and so forth.

Whether these types are generalizable to other schizophrenic populations is unknown. Beck admits that his schizophrenics are drawn from a restricted socio-

economic sampling. At present, if an independent investigator wanted to use this Q-deck with a presumably schizophrenic patient, he could sort the items and could learn whether his Q-sort correlated with one of the six types. If it did not correlate with any of the six types, then he would be left with the unsatisfactory conclusion that the person either belonged to another class of schizophrenics (not identified by Beck) or to some non-schizophrenic class. This possibility raises the question of why Beck did not use a trait universe other than one drawn from a special kind of behavior pathology. If he had, conclusions could have been drawn regarding the characteristics of his schizophrenic reaction-types as differentiated from normals and from other pathological variants.

Perhaps the most significant conclusion drawn by Beck—and reiterated time and again—is that schizophrenia is not necessarily psychosis. Schizophrenia, in any of its six varieties, is a way of life, and its ripening into psychosis depends upon the amount and kind of environmental stress.

So much for the study proper, which is a revolutionary attempt to revise psychodiagnostic theory and practice. It is too early to say whether it will have any effect on how schizophrenias are diagnosed and for what ultimate purposes. No data are provided which would allow inferences about the present or ultimate utility of the classification. Beck does suggest, impressionistically, a rank-order estimate of the degree of latency of the six reaction-types and also the degree of vulnerability to stress. The data allow only the most tentative statements about stability of the reaction-types. The discovery of empirical correlates of the types is admittedly a problem for the future.

The logic of the Q-sort in relation to the correlative Rorschach items is vague and uncertain. Did the Rorschach experts use a 'sign approach' exclusively? That is to say, did the psychologist assign a rank to a specific Q-deck item on the basis of the presence or absence of its correlative Rorschach item? Or did he operate in the more conventional mode, espoused by most Rorschach specialists, of making global judgments on the basis of the scored and unscored cues supplied by the Rorschach protocol? If the psychologist adopted the latter—more conventional—methodology, the correlative outline cluttered up the design and was unnecessary. If the psychologist operated on the basis of the molecular sign-approach, he would really have had a bad time of it. His difficulty would have been caused mainly by the vague, indefinite, sometimes double-barreled, often equivocal, characterizations of Rorschach elements. If the Q-sorts were made on this basis, the indefiniteness of these Rorschach elements would render it well-nigh impossible for another set of investigators to repeat the sorting.

The reviewer gets the impression that the monograph was hastily prepared. Besides some obvious typographical errors, there are lacunae in the presentation, and it is difficult to believe that much thought was given to the problem of organization. The writing is anything but concise (Stephenson's chapter and appendix are exceptions)—rambling to the point of distraction and often repetitious to the point of boredom.

In summary, the reviewer would assess the monograph as a somewhat disappointing report of an investigation which may be described as a pioneering effort toward establishing a more rational basis for practice and research in psychopathology.

THEODORE R. SARBIN
University of California

Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence. By PAUL E. MEEHL. Minneapolis, University of Minnesota Press, 1954. Pp. x, 149.

This excellently written little book should be *must* reading for all clinicians, would-be clinicians, and the smug statisticians. The author has maintained a judicious, fair-minded, and balanced position between the two untenable extremes represented by G. Allport and Sarbin.

The first two-thirds of the book is devoted to a so-called "theoretical analysis" of the problem. Here the general logic of clinical versus actuarial procedures is developed. The present reviewer, with his bias for the greatest possible conciseness, found some of this a bit tedious but nevertheless enlightening, especially the chapter on the logical reconstruction of clinical activity as exemplified in the theraputic setting. Apparently the therapist must pick and choose from among many hypotheses, some of which are derived from general principles (and hence have an actuarial base), others from relatively rare events, and still others from mere hunch. Seldom, if ever, can this sort of prediction be statisticized, hence the clinician as therapist need not worry about having his services replaced even in part by an inflexible multiple-regression equation. (If he is to be successful as a therapist he may need to worry about his hit-miss ratio.)

In contrast, predictions of the diagnostic-prognostic type, based on case histories, test-results, and the like, can be made by both statistical and clinical methods, and the question of which method is superior becomes an empirical problem. Accordingly, Meehl proceeds to a critical examination of 19 researches after giving adequate reasons for discarding certain studies as being irrelevant, though cited by some as relevant. His evaluations indicate that, in all but one of the 19 studies, statistical predictions are equal to or better than clinical predictions. It is, of course, not surprising to find that the clinician fails to do a better job than the statistician when both use the same basic data. What is surprising is that the clinician also fails to excel when he has access to supplementary data not available to the statistician.

Because the report of Hovey and Stauffacher is the exception among the 19 researches, Meehl says "it deserves careful study." After three pages of appraisal, he says "Out of the kindness of my heart . . . I shall score this study for the clinician." This scoring the reviewer definitely questions. Hovey and Stauffacher's hit-miss ratio of 1.7: 1 for the statistical compared with a 2.8: 1 ratio for the clinical looks impressive until converted into percentage hits: 63 and 74 respectively. It is claimed by the authors, and accepted by Meehl, that this 11-point difference is significant beyond the .01-level, but what Meehl's usually critical eye apparently overlooked is the fact that the chi-square was based on a seven-fold inflation of the basic N of 47 cases. This is the all too frequent treatment of multiple observations as though they were independent sampling units.

The title of Meehl's last chapter, "A final word: unavoidability of statistics," forecasts the conclusion expressed in his closing sentence: "Always, we might as well face it, the shadow of the statistician hovers in the background; *always* the actuary will have the final word."

Stanford University QUINN MCNEMAR

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXVIII DECEMBER, 1955 No. 4

PHRENOLOGY VERSUS PSYCHOANALYSIS

By KARL M. DALLENBACH, University of Texas

As I have pointed out on another, similar occasion,[1] controversy over theory is futile. It advertises and calls attention to the theory criticized and it is not effective against the will to believe. Besides relieving the critic's pent-up emotion, a polemical article probably accomplishes nothing; and the time, thought, and effort given to its writing are wasted. This is true because theories do not succumb to abstract argument. Criticisms are not read or, if read, are not assimilated or heeded by those wedded to the theory attacked—and they are not needed by those who are not. Theories are not killed by criticism; they do not die, they just fade away—and poor indeed is the theory of yesterday that does not find some adherents today! I have searched long through the history of our science without finding a single instance in which I could with any assurance say that criticism was the *coup de mort* of theory. Theories pass from the scientific stage not because they have been disproved but because they have been superseded—pushed off and replaced by others that are new.[2]

* Accepted for publication April 18, 1954. Presidential address of The Southern Society for Philosophy and Psychology, given at Atlanta, Georgia, April 16, 1954.

[1] K. M. Dallenbach, The place of theory in science, *Psychol Rev.,* 60, 1953, 34 f.

[2] A similar observation was made by John Dewey who wrote: "It would be difficult to find a single problem during the whole record of reflective thought which has been pursued consistently until some definite result was reached. It generally happens that just as the problem becomes defined, and the order of battle is drawn, with contestants determined on each side, the whole scene changes; interest is transferred to another phase of the question and the old problem is left apparently suspended in mid air. It is left, not because any satisfactory solution has been reached, but because interest is exhausted. Another question which seems more important has claimed attention. If one, after a generation or a century, reviews the controversy and finds that some consensus of judgment has finally been reached, he discovers that this has come about, not so much through exhaustive logical discussion as through a change in men's point of view." (John Dewey, Evolution and ethics, *Monist,* 8, 1898, 321.)

When a theory becomes the basis of a school, a *Fach,* or cult that is vigorously promoted, in which its author assumes a messianic rôle and its adherents the cloaks of disciples, in which there is an infallible pope and a hierarchy of votaries, then the critic's lot is not only futile but it is apt to be grievous as well. If his criticism is effective, if it strikes home, he is subjected to the slings and arrows of outraged disciples—seldom if ever does the pope himself deign to answer—who rise fervently to the defense of their dogma and doctrine and hide their logical lapses and *non sequiturs* in the fury of their replies. The critic may find the vehemence unpleasant but he should not be surprised by it. It is understandable; the defenders believe that their livelihood and very existence are threatened—as, indeed, would be the case if logic and experimental results alone prevailed—hence a distraction of any kind, the mere rumble of a drum, were better than the acquiescence of silence.

These sobering reflections—the futility and grief of the critic's rôle—came to mind when I considered discussing Phrenology and Psychoanalysis with you, for both are schools of the kind just mentioned. Though one is old and has, as I am sure, no defenders among you, the other is now at its zenith, with champions galore. It has more followers today than ever before and I doubt not, that, if put to the vote of the members of the American Psychological Association it would be the system chosen by the plurality if not by the majority. Though I am still vigorous enough to enjoy a good fight and am not afraid of controversy, I have no desire to play a futile rôle, hence I shall avoid criticizing the doctrines of either school. I shall restrict myself to the rôle of historian and in my review of these two schools I shall hold myself strictly to statements of fact—every one of which is well documented.

I have no thesis to defend. I merely wish to review with you the striking similarities of Phrenology and Psychoanalysis; of their invention, their development and modification, their basic philosophy and psychology, and even the actions and behavior of their *dramatis personae.*

Despite their separation in time of 100 years, they are so similar that Psychoanalysis is almost a case of history repeating itself, of new wine in old bottles; for though the voice and the words are new—the structure, form, and development of the refrain are those of Phrenology. I was struck by the similarity of these two schools some years ago when I was forced, in my search for the origin of the word 'function' as a systematic term in psychology, to make an extensive study of Phrenology,[3] and nothing

[3] Dallenbach, The history and derivation of the word 'function' as a systematic term in psychology, this Journal, 26, 1915, 473-484.

has occurred since then in the development of Psychoanalysis which leads me to reject that view. Indeed, quite to the contracy, some of the most striking similarities have appeared since the first were noted.

Few of you have any interest in Phrenology and still fewer have spent any precious time in its study. Probably the only accounts of it that many of you have seen are the shabby carcicatures that appear in many elementary textbooks of psychology as examples of a pseudoscience—of the kind of thing that psychology is not. Yet, in point of fact, Phrenology played an important rôle in the history of our science; it is the bridge between the empirical and experimental eras that is crossed by many historians who (because of the disrepute of the bridge) pay no proper recognition in toll.[4]

That my review may be easily followed, I outline it for you. I shall first give a brief résumé of the education and training of the founders (Gall and Freud) of these two schools—you will find that they are very similar; and then I shall trace in turn the development of their doctrines—you will see that history, even in minute detail, seems to be repeating itself.

(1) Founders: (a) Gall. Franz Joseph Gall, the founder-promotor of organology, the discipline now known as Phrenology, was born on March 9, 1758, at Tiefenbrunn, Baden, Germany. He was educated at Baden, at Strassburg, and at Vienna. Upon the completion of his medical training, he specialized in neural anatomy, giving special attention to the head and brain—an interest acquired in his youth from the observation that his schoolmates with prominent eyes possessed good memories. After several years of post-graduate study, he entered upon the private practice of medicine in 1785, when 27 yr. of age. While waiting for patients to come, he devoted himself during his free time to the pursuit of his hobby. He first studied the heads of people in jails and lunatic asylums, choosing these heads because their mental characteristics could be determined without question from the predicaments into which their possessors were found. He later studied the heads of friends and casts of heads of other people whose mental characteristics were known to him. In 1796, after 11 yr. of investigation and preparation, during which time he published one book,[5] he began his lectures upon organology at Vienna. They proved to be very popular; students flocked to him; he had embarked upon his career.

(b) Freud. About one hundred years after Gall's birth, Sigmund Freud, the founder-promoter of Psychoanalysis, was born on May 6, 1856, at Freiberg, Moravia, a province of Austria. When he was 4 yr. old his family moved to Vienna. He received his early training at home and in private schools. After completing his

[4] Boring is one historian who pays, but he pays with a tainted phrase for he writes that "scientific psychology was born of phrenology, out of wedlock with science" (*A History of Experimental Psychology,* 1929, 55). The kind of science this makes psychology is certainly unacceptable in a reputable family. Boring's love of turning a neat phrase trapped him as he is not the kind of man to devote a life-time at the altar of such a science.

[5] F. J. Gall, *Untersuchungen über Natur and Kunst im kranken und gesunden Zustande des Menschen,* 1791.

studies at the Wiener Gymnasium, from which he was graduated at the age of 17 yr., he decided to study medicine. It was, as he thought, the surest among the means open to him of achieving his boyhood dreams of power.[6] He entered the University of Vienna in 1873 and was graduated eight years later in 1881, his training having been prolonged by service as research assistant in the Zoölogical Experimental Station at Trieste, and also in Brücke's Biological Laboratory at the University and by a year of compulsory service in the Austrian Army as a physician. After his graduation he continued his work in neural anatomy in Brücke's laboratory for one year and then, because of his restricted material circumstances and his desire to marry, he decided to leave his beloved laboratory and to prepare himself for the practice of medicine. To that end he accepted a position in 1882 as junior physician in the Vienna General Hospital which permitted rotation among the different departments. He served two months in surgery—acquiring, as he said, "an intense dislike for it"; about seven months in internal medicine—during which he found that he had "no more interest in treating the sick patients in the wards than in studying their diseases"; five months in the psychiatric clinic—which "were highly interesting and satisfactory"; three months in dermatology—which "were welcome" because he came into contact there with syphilis, "The basis of many diseases of the nervous system," and the balance of his service, about 18 mo., in the department of nervous diseases.[7] Soon after entering this department he found himself the senior physician in charge due to the fact that all the doctors above him had volunteered to combat an epidemic of cholera that was threatening to cross the border into Austria. He enjoyed the service and the responsibility but deplored the fact that, because of the lack of knowledge concerning the cause and treatment of these diseases, so little could be done for the patients. His neurological training, which was the best his time afforded, was utterly helpless against nervous diseases. Freud resolved to do something to rectify that condition.

Hearing that Charcot, in the Salpêtrière Hospital in Paris, was successfully treating hysteria by hypnosis, Freud decided to go to him for instruction in the new therapy. He applied for and won a travelling fellowship and in the fall of 1885 when 29 yr. old he went to Paris. Charcot worked by way of post-hypnotic suggestion. He hypnotized the patients, suggested relief, and then awakened them. Though he also held the belief that there was some trouble in the sexual life of every neurotic patient, he made no reference to sex in his treatment. The idea was, however, implanted in Freud's mind and, as you all know, it germinated and later bore fruit.

Having learned the hypnotic technique, Freud returned to Vienna and started private practice as a specialist in nervous diseases. Though he continued his strictly scientific investigations and writing, more and more of his time and energy were given to clinical pursuits. He used hypnosis as his chief method of treatment. He soon ran into difficulties, however, as he found that he was not always able to induce

---

[6] Ernest Jones, *The Life and Work of Sigmund Freud*, 1, 1953, 27 f., 30, 78.

[7] It was during this period that Freud studied the physiological effect of cocaine, examples of which the Merck Company supplied him for experimental purposes. He might have discovered its anesthetic effect, which was discovered by Carl Koller, a colleague in the Hospital to whom he had demonstrated it as a "wonderful drug" that "calmed agitation and dispelled depression," but he left it slip through his fingers. (Jones, *op. cit.*, 84.)

hypnosis in his patients and, furthermore, that he could not always effect a cure by suggestion when hypnosis had been induced. To perfect his hypnotic technique, he went to Nancy, France, during the summer of 1889 to visit Bernheim, who claimed to be able to hypnotize any one. He learned there, however, that hypnosis was not necessary in therapy; that suggestion without hypnosis accomplished the same results, though perhaps more slowly.

When he returned to Vienna after the Nancy pilgrimage, he joined forces with Breuer, an old friend, and tried the 'talking out' method that Breuer had used in a case of hysteria, the famous case of Anna O., during the early 1880s. They obtained, as they thought, excellent results and jointly published a 14-page paper in 1893 on "The psychic mechanisms of hysterical phenomena" and, in 1895, a 269-page book entitled *Studies on Hysteria.* With the publication of this book, Psychoanalysis was launched.

The founder-promoters of our two systems, Gall and Freud, were both medically educated—they received the best training that their different generations afforded. They both specialized in neurology and each made outstanding contributions in that field. Their training was not hurried—they both served long apprenticeships. They were able, brilliant, and talented; men possessed of great zeal, determination, and imagination; destined to carve their niches in history! Now let us see what niches they carved.

Systems: (a) Phrenology. Gall based his system upon the most advanced physiology and psychology of his day.[8] He held that the brain was the organ of mind—a doctrine advanced by the Greeks and repeated with variations by many authors since then but never stated by anyone before him with such clarity and finality. Not some but *all* the various aspects of mind were located by Gall in the brain. His psychology was also the most modern of his time. Of the schools then in existence, associationism and faculty psychology were the most prominent. Of these Gall chose faculty psychology. That it was soon to receive its death blow at the hands of Herbart, Gall could not know, but had he chosen associationism he would not have fared any better. Faculty psychology, however, suited his system; it was respectable during his time and he had no reason to question his choice. Indeed, his adoption of faculty psychology and not Herbart's criticisms may have hastened its passing.

Gall correlated the various faculties, into which he divided mind, with parts or organs of the brain—the amount of a given faculty being related to the size

[8] For an excellent account of Gall's intellectual antecedents see Madison Bentley, The psychological antecedents of phrenology, *Psychol. Monog.*, 21, 1916 (No. 92), 102-115.

of its corresponding cerebral organ. Assuming that the skull is molded in its shape by the growing brain, he believed that the outer surface of the skull and the contour of the brain were sufficiently close that knowledge of the relatives sizes of the organs of the brain and the corresponding faculties could be obtained by examining the surface of the head. Specifically, an enlargement at a spot on the skull was taken to mean a corresponding enlargement of the brain and an excess of the faculty located there; contrariwise, a recess or indentation in the skull meant a deficiency in the brain and a lack of the corresponding faculty.

Gall's method of investigation was empirical. The experimental method was still unknown in psychology. It was years away in the future—about 40 before Weber reported his experiments upon touch, 60 before Fechner published his psychophysical investigations, and 80 before Wundt presented his program. Gall used the only method that was then at hand—a rigorous search for instances illustrating his doctrine. He labored assiduously to multiply the number of his observations. He examined the heads of people in every station and rank of society—of men of talent and of genius, of the mighty and of the humble, of criminals, imbeciles, and insane, and even of animals. He compared successive generations for common traits. For all his pains and labors, his system was, however, an invention, not a discovery. Instead of deducing the faculties from the organs and generalizing both from specific observations, Gall selected the faculties and searched for positions on the skull at which to assign them. His localizations were frequently based upon striking instances. For example, he placed *cautiousness* in the parietal area because an ecclesiastic with a hesitating mien had large protuberances there, and he placed the *love of approbation* at the top and toward the rear of the head because a lunatic claiming to be the Queen of France had large bulges there. He sought to substantiate the localizations made from striking cases but his search was always directed toward obtaining confirmatory evidence.

Gall did not hesitate, however, to modify his localization in the light of new observations (when made by him) or criticisms (when he deemed them pertinent—which he seldom did). The first edition of his craniology, published in 1800, contained 22 duplicated organs which were arranged at different levels of the two halves of the brain: some at the base, some about the center, and others on the surface. The organ of *vitality,* the desire to live, was, for example, localized in the medulla oblongata and measured by the size of the foramen magnum and the thickness of the neck. When asked: "How do you know that the prominence of an outer organ indicates its real size? May it not merely be pressed out, though itself of inferior volume, by the development of a lower organ?"—Gall had no answer; or rather his answer was immediately to withdraw the separate organs at the base and about the center of the brain and to extend them all the way up from the base to the cortex—thus avoiding that criticism. In successive editions of his charts, the number, names, locations, sizes, and shapes of the organs were changed, but in all of them the organs were represented as round or oval enclosures with vacant interspaces. By 1807 the number of organs had been increased from 22 to 27.

Though Gall's doctrines were rejected by his colleagues in Vienna, they were of a kind to find ready converts among the laity. His lectures had tremendous popular appeal. Here was a short cut, a key to the mystery of personality and the self for which the world was looking—and for that matter still is! In 1802, at the height of his popularity in Vienna, his lectures were interdicted by the Government as

being dangerous to religion. This apparent calamity proved a boon, for he and Spurzheim, who was now associated with him as a collaborator, carried their investigations and lectures throughout Europe, exchanging the audience of a city for that of a continent. They settled in Paris in 1807 where their lectures and researches were well received by the public and, in contradistinction to their treatment in Vienna, were respectfully considered by their scientific colleagues. In 1808 they were invited to present a *mémoire* upon their work to the Institut de France. Though such invitations were usually tantamount to election, this did not follow in their case. The Committee, to which their *mémoire* was referred for consideration, reported (at the suggestion of Napoleon, it is believed) that the thesis presented did not fall within the scope of its field (mathematics and physics), which it certainly did not. Napoleon, as the apocryphal story goes, was displeased by the recent election of an Englishman and decided against the election of other foreigners. As founder and patron of the Institute, Napoleon's recommendations would, of course, have been decisive.

Undaunted by this rejection but still pleased by the recognition they had received, Gall and Spurzheim continued their researches and published the first two volumes, in 1810 and 1812, respectively, of an anticipated four-volume work. By this time, however, their relations were beginning to be strained. Spurzheim—student, assistant, disciple—was beginning to feel his maturity. His co-authorship of the *magnum opus* of the new system, in particular, led him to believe that he was of equal stature with the master—which he was as a lecturer and promoter—and he pressed for changes that Gall thought were radical. He wished to add to the number of faculties and organs, which had remained static for many years; to systematize the terminology, which was badly in need of it; and so to change the shapes of the organs on their charts that there should be no vacant interspaces. Gall, the founder and master, would have none of this, with the inevitable result that the long and close association was broken. In 1813 they went their separate ways. Gall remained in Paris to continue his research, writing, lecturing, and practicing, while Spurzheim fared forth to spread the gospel. Soon after the break, whether to make it complete or to differentiate his teachings from those of Gall, Spurzheim adopted the term 'Phrenology,' which was coined by Thomas Foster in 1815 from the Greek word meaning 'mind.' Though Gall adhered to his term 'organology' and wrote and spoke of its different aspects as 'physiognomy' and 'craniology,' Foster's term prevailed, due probably to the greater, world-wide publicity given it by Spurzheim.

Spurzheim traveled extensively through Germany, Switzerland, England, Scotland, and Ireland, and he died in Boston in 1832 while presenting the doctrines of phrenology in America. He was a fluent lecturer, a convincing demonstrator, but withal a promoter, a protagonist, rather than a scientist. He was too busy disseminating the doctrines of phrenology to engage in research. For the scientific and theoretical support of his doctrines he continued to draw upon Gall even after their break. His skill in controversy and his engaging personality won Phrenology many converts. For example, in 1815, arriving in Edinburgh, Scotland, after an article on phrenology in the *Edinburgh Review* had been roundly denounced, he gave a series of lectures and demonstrations, which so convincingly refuted the critics that for many years this city was a phrenological stronghold. He made many friends and adherents for phrenology during this visit, foremost among whom were the Combe brothers, George and Andrew.

George Combe, a strong candidate for the then vacant chair in logic at the University of Edinburgh, had derided and strongly opposed phrenology. After Spurzheim's visit, however, he took up the cause and made it his own until his death in 1858. His conversion was complete; he was the Scottish Paul of the phrenological gospel. Combe published his first article on phrenology in 1817, his first book in 1819, he founded the Phrenological Society of Edinburgh in 1820, and *The British Phrenological Journal* in 1823. He wrote and lectured extensively on phrenology and, like Spurzheim, he carried the word to America.

Phrenology spread rapidly in America, the land of the free and, according to Barnum, the home of people who like being humbugged. The Fowler brothers were the first and chief supporters of the cause. They wrote many books and articles on phrenology and founded the New York Institute of Phrenology and *The American Phrenological Journal* in 1838. Institutes at which the science and art of phrenology could be learned were established in all the larger cities and soon small indeed was the city or town that did not have its own practicing phrenologist.

The spread in conservative England was as phenomonal. By 1832 there were 29 phrenological societies there and numerous phrenological journals, the first of which, *The British Phrenological Journal,* as already mentioned, was established in 1823.

During this period of rapid territorial expansion, the number of powers and organs also increased. Before his death, Spurzheim had raised the number to 37, all being located upon the skull in contiguous patches. Later phrenologists extended the number to 39 and to 41, and a few to 43. How right Gall was! Once the process of multiplying was begun there was no stopping it within the ingenuity of man to devise powers and to find places on the skull and brain for the correlated organs.

The doctrines of phrenology were never generally accepted by men of science. Most of them, after a cursory examination, which sufficed to disclose the inadequacies, ignored phrenology and went about their own proper business. A few however, like Sir William Hamilton, for example, unwisely decided to criticize phrenology and to point out its pitfalls—I say "unwisely," because, as previously remarked, a doctrine does not yield to criticism and their efforts were poorly rewarded; they lost precious time from their own work and did not save others with the will to believe from the pitfalls they had so well and carefully marked. Phrenology endured as a controversial doctrine nearly 40 yr. after Hamilton had written his last critical paper. It was the growth of knowledge, not criticism, that pushed phrenology from the scientific stage. In addition to large numbers of the general public, a small group of scientists accepted phrenology's doctrines, and some members of this small group —Herbert Spencer and August Comte, for example—were men of first importance. They were the jewels in phrenology's crown but Spencer retracted in 1855,[9] hence his luster, if not that of Comte, was only temporarily dimmed.

Gall was grieved and bitterly disappointed by the reception of his doctrines, in particular by his confréres of the medical profession who either completely ignored or roundly condemned them. When he realized that hope for their approval was vain, he retaliated by divorcing medicine from his house—maintaining that a phrenologist did not need a medical education and that the proper way of becoming a phrenologist was to study phrenology. With this pronouncement, phrenology passed

[9] Herbert Spencer, *The Principles of Psychology,* 1855; American ed., 1, 1870, (§248), 572-576.

into the hands of the laity—at first men of good training and education: teachers, lawyers, ministers; but as the practice of the phrenologist became more lucrative and institutes of phrenology more numerous, the caliber of the students and of the practitioners fell.

Phrenology enjoyed its greatest popularity about one hundred years ago in the 1850s. From then it began to wane, slowly at first, then more and more rapidly. Its decline was not caused by criticism but by the development of new interests, of a *new phrenology,* of the experimental results of the brain physiologists: Broca, Fritch, Hitzig, Ferrier, Goltz, Munk, and a host of lesser lights.

Phrenology, however, lingered on. Its advocates struggled against the inevitable. The department of psychology in one of our large eastern universities was founded in the late 1880s by a wealthy devoteé of phrenology in the hope and expectation that the studies and writings of its staff would advance the doctrines of phrenology. In 1908 it was still profitable to reprint Spurzheim's book on *Phrenology or the Doctrine of the Mental Phenomena,* which was first published in 1825. *The American Phrenological Journal* ceased publication in its 124th volume in 1911, and the American Institute of Phrenology was in existence as late as 1925. In many of our larger cities today, especially on the West Coast, phrenological practitioners may still be found. The invention, rise, and decline of phrenology occupied a little over a century. Gall's niche in history, never very large, is now but a roost for charlatans.

(b) Psychoanalysis. Now let us return to psychoanalysis, which we left at its launching in 1895, and follow its course until the present. Except for a shift in dates of 100 years, a change in the names of the actors and their parts, the drama of psychoanalysis is similar enough to be a plagiarism of the one just told.

Studies on Hysteria, published in 1895, as earlier noted, was nothing more than an account of a new therapy—"a new method of treating and curing hysteria." The theoretical assumptions behind the studies were two in number: (1) hysterical behavior is a substitute for normal psychic acts and it possesses, therefore, meaning and significance; (2) when this meaning stands revealed to the patient, the troublesome symptoms disappear. The *Studies* gave a certain amount of plausibility to these assumptions, but more proof was needed. Freud was left alone to supply it when Breuer withdrew from collaboration with the publication of this book.

Various explanations are given for Breuer's withdrawal. Probably the true reason was his unwillingness to undergo the embarrassment of the phenomenon known as 'transference' that is frequently encountered in the therapy, *i.e.*, an emotional involvement of the patient with the physician. Breuer's famous case of Anna O. did not, it seems, end so happily, as the account given in the *Studies* indicated: On the basis of

information received years later from Freud by word of mouth, Jones reveals that after Anna O. had been pronounced cured, she had a relapse (a pseudocyesis) and that Breuer then had considerable embarrassment in terminating the treatment.[10] To quiet his patient at her last visit, he "hypnotized her, fled the house in a cold sweat, and the next day he and his wife left for Venice" for a long vacation. Confirmation of this story appears in one of Freud's letters to his wife in which he wrote (in reply to her anxiety that he too might find himself in a similar embarrassing situation) that she need have no fear, "for that to happen one has to be a Breuer."[11]

Whatever the reason for Breuer's defection, he withdrew and Freud was left alone. The book, *Studies on Hysteria*, was not well received by the medical world. The reviews were few and all were unfavorable. Of the 800 copies printed, not many were sold. Except for the popular appeal of Phrenology, which the *Studies* did not have, Freud's situation was the same as Gall's 100 yr. before. Freud, however, did not at that time have a system to defend—he had merely a therapy to save. To accomplish this end he sought to accumulate cases, as Gall had done. Since the method was long and tedious, he sought ways of improving and of shortening it. He dropped hypnosis but continued to place his patients in a reclining and comfortable position. He tried various procedures: he sat where they could see him, or at their heads where they could not; he gently stroked their foreheads, or refrained from touching them. He encouraged them to talk, talk, talk; now directing their monologues, now permitting free association. He seized upon complex indicators: hesitation, avoidance, reports of trivialities, evidence of embarrassment; and directed the 'talk' upon these. It was still slow work, requiring daily sessions for months. He hunted therefore for a short cut—for some means of catching the 'unconscious' off guard—that would disclose the repressed emotion and reduce the expenditure of time.

He turned first to the dream, upon which he published two books: a large one of 510 pages called *The Interpretation of Dreams* in 1899, and a short one of 110 pages entitled *On Dreams* in 1901. He turned, secondly, to the slips of the pen, tongue, and memory which led to the publication in 1904 of *The Psychopathology of Every Day Life*; and, 'since many a true word is spoken in jest,' he turned, thirdly, to the play of wit and humor which resulted in the publication of *Wit and Its Relation to the Unconscious* in 1905. In the same year he published *Three Contributions to the Theory of Sex* in which Charcot's suggestion that there "was some trouble in the sexual life of every neurotic patient" came to its full fruition. While producing and publishing these books, Freud continued to write numerous short articles and case histories illustrating the method and efficacy of his therapy.

Until the appearance of these books, Psychoanalysis was only a technique for treating neuroses; but now, by some strange magic, difficult to comprehend, the books, by their mere existence, transformed the therapy, in Freud's mind at least, into a system of psychology which stood in opposition to all others. This Gargantuan step was taken despite the fact, be it remembered, that these books were merely the results of attempts to find short cuts in the therapy, and that the conclusion—that the short cuts worked—was a matter of opinion and not a result of experimentation! Freud failed to see, as Gall failed before him, that examples merely illustrate but never prove; that opinions set problems but never solve them.

[10] Jones, *op. cit.*, 224. [11] *Idem*, 225.

Freud's method of demonstrating his propositions was, like Gall's, empirical. Though the experimental method was at hand, as it was *not* in Gall's day, Freud made no use of it. He disliked its "tedious exactitude."[12] Jones writes in his biography of Freud, that Freud essayed the experimental method three times and each time unsuccessfully,[13] and that the only experimental study that Freud ever published is of interest because "its rather dilettante presentation shows that this was not his real field."[14] He never undertook a study in which the tenets of scientific procedure—repetition, variation, isolation—were observed. He had no contact with the psychology of his day and generation. Regarding his knowledge of psychology, Jones has this to say: "Freud was . . . ill-informed in the field of contemporary psychology and seems to have derived only from hearsay any knowledge he had of it. He often admitted his ignorance of it, and even when he tried to remedy it later did not find anything very useful for his purpose in it."[15] Not being schooled in psychology, he used its terms carelessly and inaccurately—for instance, 'perception,' 'sensation,' and 'idea' were used interchangeably. He coined his own psychological terminology.

Like Gall, he never subjected the data of any of his studies to statistical treatment. Though the statistical methods were at hand, as they were *not* in Gall's day, and their use was commonplace, Freud was completely uninformed in regard to them. This is not strange as his scientific training and competency were in the field of neurology in which the statistical treatment of data was of little or no importance. In tracing neural pathways, one microscopic slide showing a connection was sufficient, more were not needed. He was, furthermore, like Gall, interested in improving the lot of man. He was not a 'pure' scientist; he was an artist, a technician, his aim was to devise a method for treating neuroses.

With the appearance of his books, Freud began to gather followers. Not being a member of the Vienna faculty, he met and instructed them in his home. From a small local group of inquirers, the circle widened; students came from other countries, chiefly Switzerland, and by 1908 the number was sufficiently large to dignify the 'new' discipline by the convocation of an International Congress of Psychoanalysis.

In 1909, G. Stanley Hall, innovator, eclectic, and a man of catholic interests, invited Freud to speak upon his doctrines at the celebration of the twentieth anniversary of the founding of Clark University. This was an acolade to Freud—similar to the invitation extended Gall and Spurzheim in 1808 to present a *mémoire* to the Institut de France. Freud said in regard to this invitation, the first public recognition he had received, that it "encouraged my self-respect in every way. In Europe I felt as though I were despised; but over there [in America] I found myself received by the foremost men as an equal."[16] The papers read by Freud at the Clark Symposium were published in *The American Journal of Psychology*[17] and the doctrines which had been ignored by academic psychology as unclear and untouchable were introduced to the psychological world and given the cloak of respectability.

When Freud returned home, he set about organizing the International Psychoanalytic Association and, at its first meeting in Vienna in 1910, he put "the leadership . . . into the hands of the Swiss" and moved Jung's election as the first presi-

[12] Jones, *op. cit.*, 40. [13] Jones, *op. cit.*, 54.
[14] *Idem*, 92. [15] *Idem*, 371.
[16] H. A. Murray, Sigmund Freud: 1856–1939, this JOURNAL, 53, 1940, 135.
[17] Vol. 21, 1910, 181-218.

dent, because, foreseeing the future, he thought the Swiss might save the movement.[18] Freud, however, remained its supreme arbiter.

The Association was a strange organization. It was more like a militant, religious cult than a scientific society, and in that respect it was very like the Phrenological Societies promoted by Spurzheim. The International Association established branches in various countries and publishing houses and periodicals for the promulgation of its doctrines; it arranged for the translation of Freud's books; it determined its membership upon the basis of orthodoxy; and it extended the psychoanalytic method and principles into every phase of human life—into religion, law, literature, anthropology, myth, and custom. Nothing human was foreign to its probing.

So enthusiastic were the members of the new organization that some of them surpassed Freud in his extension of the doctrines. This could not, however, be tolerated by Freud, hence the International organization had hardly been formed before it began to break up. Adler, Jung, Rank, Stekel, and others were unable to accept pan-sexualism as the central doctrine of psychoanalysis and they argued against it and brought facts of observation to bear in its disproof, with the consequence that Freud, in 1912, moved their expulsion from the Association. A break between master and disciples, such as occurred in 1813 in Phrenology! Freud offered no apology for his actions. He simply said, in reference to them, that "Psychoanalysis is my creation and I feel myself justified in assuming that nobody knows better than I what psychoanalysis is."[19] He explained his dogmatism upon the grounds that it was necessary to preserve his doctrines from disintegration and dilution. "To permit heresy within the ranks of the newly established school would be to invite disaster. Since no university had dedicated a chair to the newly created science and there were no legal formalities . . . to prevent any charlatan from setting up an office and practicing psychoanalysis, it was the duty and obligation of the International Psychoanalytic Association to perform the dual function of instruction and certification." Freud's aversion to 'scientific' polemics went so far that he later "urged that psychoanalytical congresses should be confined to the reading of papers—followed by reflection, testing, and perhaps private discussion."[20]

The war years of 1914–1918, with their shell shock, fear, and neurosis, accelerated the spread of the psychoanalytical doctrines. To meet the need for information regarding the subject, Freud published, in 1916, *A General Introduction to Psychoanalysis* which he intended to serve as a textbook. It was the best single source regarding Freud's teaching published up to that time.

The postwar years were fruitful ones for Freud. In addition to numerous articles, he published a book a year for the next five years. In 1923 he published *The Ego and the Id* in which he explained the super-ego. Freud added here three faculties to those already in his system, with the same aplomb that Spurzheim exhibited when he added other faculties to Phrenology.

Though Freud's followers were to quote an enthusiastic disciple, "as grasses of the earth . . . who held his books as scripture and him in idolatrous regard," his doctrines by 1925 had not received the attention and recognition that he thought they deserved. Despite their introduction to the psychological world through *The*

---

[18] Fritz Wittels, Revision of a biography, this JOURNAL, 45, 1933, 748.
[19] Murray, *op. cit.*, 135.
[20] Jones, *op. cit.*, 257.

American Journal of Psychology, the doctrines were, for the most part, completely ignored by the academicians. Wundt, Müller, Titchener, Külpe—call the roll and you will find that the large majority of the great in psychology during the first quarter of this century ignored Freud, and that the few who did not were highly critical of his doctrines.[21] Though Meumann,[22] in 1907, and Köhler,[23] in 1912, wrote on dreams (Köhler extensively), neither referred to any of Freud's writings upon that subject. Külpe, writing on Psychology and Medicine in 1912, did not mention Freud's therapy or even his name.[24] Freud's doctrines were anathemas to the academicians during this period, even as Gall's were to the academicians of his day. I well recall the first time Freud's name was mentioned in a meeting of the American Psychological Association. It occurred during the Christmas meetings in 1923 at the University of Wisconsin. During the discussion following the reading of a paper—the kind of thing that Freud had forbidden in his meetings—a member of the audience started to tell how Freud would explain the results obtained. Before he had proceeded far, J. McKeen Cattell arose and, after expressing astonishment and painful suprise that a member of the Association should be so wanting in wisdom as to introduce Freud's name at a scientific meeting, castigated him for his folly, as only Cattell could do.

Freud did not, however, much mind being ignored by the psychologists for he had in his turn ignored them, but he was sorely grieved by the neglect of the medical profession. In 1923 he developed cancer of the jaw. Expecting an early death and despairing of gaining the support of his confrères during his lifetime, he turned his back upon the medical profession, as Gall did one hundred years before him. In 1926 he published his book upon *The Problem of the Lay-Analyses* in which he asserted that a medical education was not only unnecessary for an analyst but that it was positively detrimental. It was unnecessary because it included many subjects that the analyst did not need—chemistry, anatomy, physiology—and it was detrimental because is predisposed the student to a materialistic therapy and blinded him to the psychical method of treatment. A medical education was, more-

---

[21] Among his critics from American psychology are: Knight Dunlap, *Mysticism, Freudianism, and Scientific Psychology,* 1920, 1-173; *Old and New Viewpoints in Psychology,* 1923, 62-67; Joseph Jastrow, *The House That Freud Built,* 1932, 1-295; William McDougall, Professor Freud's group psychology and his theory of suggestion, *Brit. J. Med. Psychol.,* 5, 1925, 14-28; A great advance of the Freudian Psychology, *J. Abn. Psychol. & Soc. Psychol.,* 20, 1925, 43-47; *Psychoanalysis and Social Psychology,* 1935, 1-325; R. S. Woodworth, Some criticisms of the Freudian psychology, *J. Abnor. Psychol.,* 12, 1917, 174-194. E. L. Thorndike, in a review of Freud's *New Introductory Lectures on Psychoanalysis,* wrote that "Freudianism is like phrenology, productive of facts which would have been discovered in any case and to put to better use without it." (*New York Times, Book Review Section,* November 26, 1933, 4.) William Stern, in Germany (Die Anwendung der Psychoanalyse auf Kindheit und Jugend: Ein Protest, *Zsch. f. angew. Psychol.,* 8, 1914, 71-101), and F. H. Bartlett, in England (The limitations of Freud, *Sci. & Soc.,* 3, 1939, 64-105), gave serious consideration to Freud's system and found it wanting.

[22] Ernst Meumann, Ueber Organempfindungsträume und eine merkwürdige Traumerinnerung, *Arch. f. d. ges. Psychol.* 9, 1907, 63-70.

[23] Paul Köhler, Beiträge zur systematischen Traumbeobachtung, *ibid.,* 23, 1912, 415-486.

[24] Oswald Külpe, Psychologie und Medizin. *Zsch. f. Pathopsychol., Monog. Suppl.,* 1, 1912, 1-81.

over, inadequate, because it omitted many subjects from the curriculum that were essential to the training of an analyst—ethnology, mythology, religion, folk-lore, and the psychology of the unconscious. What was needed most of all, however, was for the student himself to be psychoanalyzed that he might be freed from his own repressions before he undertook to free his patients from theirs.

In the year following the appearance of this book (1927), the International Psychoanalytic Association debated the question: Should the Association approve and train non-medical analysts? Freud argued the affirmative and closed with the statement that he was not certain whether the wish of the doctors to possess psychoanalysis as a monopoly was due to a desire to preserve or to destroy it. In any case, such a monopoly was, as he said, impractical and equivalent to repression. The affirmative was sustained and psychoanalysis, freed from the restrictions of medical training, entered upon a lusty period of growth. Despite his cancer—which fortunately yielded to surgical treatment—Freud continued at the helm of his system for 12 more busy, productive, and dramatic years. He died in London, a refugee from Nazi-Germany, on September 23, 1939, at the age of 83 yr.

Despite his misgivings, Freud lived to see his doctrines accepted by a large proportion of the medical profession concerned with the treatment of nervous and mental diseases. At his death, with his guiding hand removed, Psychoanalysis began to splinter—even as Phrenology did after Gall's death. We have in consequence almost as many analytic schools now as there are teachers. Students are now discouraged from reading Freud. They are assigned more recent texts, written by their teachers, which bring Psychoanalysis up to date. Since knowledge is not static, this would be a healthy situation if only there were a way of determining whether the changes of doctrine were advances; but Psychoanalysis has no better method of demonstrating improvements than Phrenology had. Yet, despite disagreements, which arose among Freud's disciples as soon as the master was removed, the popularity of the analytic school continued to grow. Interest, as you all know, mushroomed during the traumatic years of World War II and it is probably greater now among psychiatrists and clinical psychologists than at any other time in history. Phrenology, as you will recall, enjoyed its greatest popularity in the 50s of the last century and then faded away. Will the same fate befall Psychoanalysis? Will our intellectual descendants a hundred years hence look back upon the followers of Freud with the same amused feelings of superiority that we enjoy as we look back upon those of Gall?

How can psychoanalysis be wrong when so many people, some of whom rank high in our science, accept it? Numbers are of no weight, and even men of prominence may be wrong, which Phrenology so well attests. As many, proportionately, 'believed' in Phrenology in its day as believe in Psychoanalysis today, and the prominence of the 'believers' is similar in

both cases. Why should clinical psychologists today accept what their confrères of 30 years ago either ignored or denied? I do not know, I wish I did; my guess, however, is that it is due to their lack of training in philosophy. The older psychologists usually studied philosophy; either elected it as a minor or took several courses in it during their graduate or undergraduate training. The psychologists of today have little acquaintance with any theory of knowledge. The older generation argued the mind-body problem; it knew the logical consequences of the assumption of powers and faculties; the present generation does not worry over such obtuse problems. A paper as long as this one could be written upon the topic, "The value of the study of philosophy to the psychologist." I suggest this as a topic to our next president who is to be a philosopher.

Now for one more question: What does the future hold for Psychoanalysis? In the light of our knowledge of the past, I could make a prediction but I prefer to let Freud speak for his own therapy. He said, according to Ernest Jones, his biographer, "that in time to come it should be possible to cure hysteria and nervous diseases by administering a chemical drug without any psychological treatment."[25] If that should come to pass, as many believe it will (and as we have some evidence today), what will that do to Psychoanalysis? What then will be Freud's niche in history? Will it, like Gall's, became a roost for charlatans? Is is not already occupied by some? I leave these questions with you.

---

[25] Jones, *op. cit.*, 259.

SHIFTS IN BINAURAL LOCALIZATION AFTER PROLONGED EXPOSURES TO ATYPICAL COMBINATIONS OF STIMULI

By Richard Held, Brandeis University

Following Berkeley's example,[1] many authors have attributed to previous experience part or all of the ability of the human listener to localize sounds in space.[2] The literature on sound-localization, though rich in studies of the stimulus-cues, contains little or no experimental evidence to support this point of view. Casual observations that children can localize correctly throughout the period of growth of their heads suggests, however, that experience is involved at least in the maintenance, if not in the actual acquisition, of accurate localization. The direction of the localizing response made by a listener to sound has been shown to depend in part on the small differences between the acoustic stimuli that arrive at his two ears. The magnitude of these differences, which are called the binaural cues,[3] increases systematically with the angular displacement of the direction of a sound-source from the median (sagittal) plane of the head of the listener. In addition—and crucial for present purposes—if the dimensions of the listener's head increase, as they do in childhood, so do the magnitudes of the binaural cues produced by a source in any constant direction from the head except the median plane.

From birth until completion of bodily growth, the linear dimensions of the head increase to one and one-half times their original sizes.[4] Consequently, from the beginning of accurate localization, which is reported to occur within a few weeks after birth, the direction of response made to almost any binaural cue must shift during growth. Otherwise systematic errors of direction would be found at some time during this

---

* Accepted for publication October 22, 1954. This study, made at the Psycho-Acoustic Laboratory of Harvard University, was conducted under Contract N5ori-76 between Harvard University and the U. S. Office of Naval Research U. S. Navy (Project NR142-201, Report PNR-141).

[1] George Berkeley, *An Essay towards a New Theory of Vision,* 1709, sects. 46-47.

[2] E. G. Boring, *Sensation and Perception in the History of Experimental Psychology,* 1942, 381-383.

[3] The history and meaning of this notion is discussed in Boring, *op. cit.,* 381-392, 397-398. A summary of present information on binaural localization is contained in J. C. R. Licklider, Basic correlates of the auditory stimulus, in S. S. Stevens (ed.), *Handbook of Experimental Psychology,* 1951, 1026-1030. A bibliography in this field can be found in Psycho-Acoustic Laboratory, *A Bibliography in Audition,* 1950, 2 vols.

[4] A. Meyer zum Gottesberge, Physiologisch-anatomische Elemente der Schallrichtungsbestimmung, *Arch. f. Ohrenk.,* 147, 1940, 219-249.

period. To account for this shift during childhood, we may suppose that recurrent stimulation during exposure to the normal environment in some way conditions the localizing response. Appropriately changed conditions of exposure may produce similar shifts of localization in adult listeners. The experiments to be described concern the effects of exposure to sounds upon a mobile observer whose binaural axis was displaced by a small angle around the vertical axis of his head by means of a pseudophone. The effects were measured by comparing the directions of response made under comparable conditions before, during, and after periods of exposure. Two experiments are reported: Experiment I, exposure to a normal environment; and Experiment II, exposure under controlled conditions.

Apparatus and Procedure

Pseudophone. This device consisted of two matched hearing aids that substituted microphones for the outer ears as receivers of air-borne sound. Its purpose was to effect a rotational displacement of the outer ears around the vertical axis of the head and at the same time to attenuate sounds reaching the ears directly. Each hearing aid consisted of a flat crystal microphone 1-in. square that fed into a battery-powered amplifier whose output drove a minature earphone. The earphone was fitted with a hollowed tip that was inserted into *O*'s auditory canals. Fig. 1 shows *O* wearing the pseudophone. The two microphones, M_R and M_L, with front surfaces in the same plane facing upwards, were shock-mounted 20 cm. apart on the ends of a rod called the microphone axis m–m. A short shaft fixed at right angles to the center of this rod fitted into a bearing B set on top of the headband worn by *O*. The center of the microphone axis was approximately 20 cm. above the center of the interaural axis i–i. By rotating the shaft, the axis of the microphones could be set either parallel to the interaural axis,[5] as shown in Fig. 1, an approximation to the normal receiving positions of the two ears, or displaced from it by 22° clockwise or counter-clockwise around the vertical axis a–a of the head. The headband and ear cushions that fitted over both the ears and the miniature earphones fixed the positions of the microphones relative to the head. External sounds (unfiltered clicks) that entered the ear directly were attenuated by approximately 40 db. *O* also wore a light jacket, shown in Fig. 1, which contained the amplifiers and batteries of the hearing aid. The entire apparatus restricted *O*'s movements only slightly.

Measuring device and procedure. Two considerations influenced the design of the measuring apparatus: (1) the greatest accuracy of directional discrimination is obtained when the source to be localized lies in or near the median plane of *O*'s head, thereby producing time-differences close to zero; and (2) ambiguity of localization (front-back reversal) is eliminated when *O* is allowed to rotate his head in the presence of the source.[6]

---

[5] The interaural axis was assumed to be approximated by a straight line intersecting the two tragi.

[6] Hans Wallach, The role of head movements, and vestibular and visual cues in sound localization, *J. Exper. Psychol.*, 27, 1940, 339-368; Herbert Klensch, Beitrag zur Frage der Lokalisation des Schalles im Raum, *Arch. f. d. ges. Physiol.*, 250, 1948, 492-500, 706-713.

For the measurements, O was seated in a chair that he could rotate by means of leg movements around the vertical axis a–a continuous with that of his trunk and erect head. The microphone rod m–m was clamped at right angles to a vertical shaft, S, lying along the same vertical axis a–a. The angular position of this shaft was indicated by an attached pointer that moved around a circular dial, D, above O's

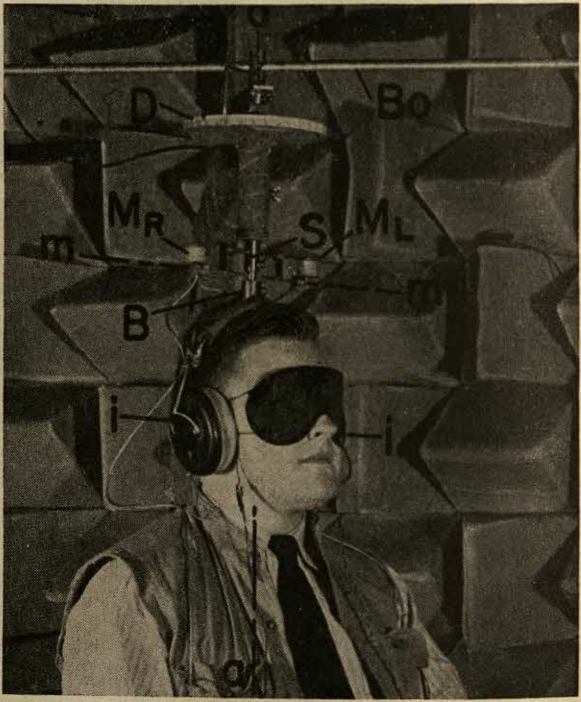


FIG. 1. O WEARING PSEUDOPHONE IN ANECHOIC ROOM

M_R and M_L, right and left microphones; m-m, microphone-axis; bearing B allows rotation of microphone-axis around vertical axis of O's head. Parts of the measuring device include: S, vertical shaft to which pseudophone is clamped; D, dial indicating rotation of vertical shaft and direction of sound-source; Bo, boom carrying sound-source. i–i is O's interaural axis; a–a is the vertical axis of rotation.

head. Above the dial, a horizontal boom, Bo, was mounted in a bearing that was centered on the axis a–a. The boom carried a sound-source (Permoflux PDR-8 earphone) 1-m. distant from the axis a–a and in the horizontal plane defined by rotation of the microphone axis. The angular position of this source—the direction of the boom—could be read from the same dial, D. Directional localization with the unaided ears could be measured with this apparatus by removing the pseudophone and installing a biting board. An L-shaped rod, clamped to the shaft S, held the biting board suspended at O's mouth. The angular orientation of the biting board could also be read from dial D. The rigid connections between microphones and shaft or between biting board and shaft ensured that the identical measuring conditions occurred after removal and replacement of O in the measuring apparatus.

All measurements were taken in an anechoic room. The experimeter fixed the source at an angle to *O*'s median plane. *O*, who was blindfolded, was then instructed to turn himself back and forth until the source appeared to be straight ahead of him. He was requested to orient his head, trunk, and even eyes in the direction of the source before indicating completion of his judgment. *E* then recorded the angular direction of the source from *O*'s head. To ensure the independence of successive judgments, the source was displaced alternately to the right and left of *O*'s median plane by angles that varied from 25 to 60°. None of the *O*s confused the true frontal direction of the source with the 'phantom' rearward direction. So consistent were the localizations, that sets of 10 to 20 successive measurements taken under optimal conditions gave standard deviations between 2 and 3°. All judgments were taken with the axis of the microphone normalized *i.e.* set parallel to *O*'s interaural axis.

Acoustical stimuli. The sounds reaching *O*'s ears, when he wore the pseudophone, were determined by the following conditions of transmission and transduction.

(1) Sounds arriving at the microphones. The energy of the sounds arriving at the microphones of the pseudophone depended upon (a) the distance and direction of the source, (b) the sound emitted by the source, and (c) the transmission characteristics of the space containing the source and the microphones. The mounting of the microphones ensured that the sounds activating each of the microphones did not differ appreciably in intensity regardless of the direction of the source and the frequencies emitted. Furthermore, since all the measuring and the entire procedure of Experiment II were carried out in an anechoic room, reverberation was not significant.

(2) Transmission characteristics. With the sound-source set in the median plane of the pseudophone, the sound-pressure outputs at both earphones were made practically identical for frequencies up to 3500~, and over the intensity range of sounds maintained under experimental control. Above 3500~ small inequalities, not in excess of 6 db., were present. The gains of both channels were kept equated within 2 db. throughout the experiments, as shown by periodic checks of the voltages across the earphones. The phase relations of the two channels were such that the diaphragms of the two earphones moved simultaneously either towards or away from their respective eardrums.

To eliminate the inequality of response above 3500~, the sound-stimuli used throughout Experiment II and in all measuring were restricted to a frequency band between approximately 1000 and 2450~. Electronic pulses at 60~ were fed through a mel-band filter having an attenuation of 50 db. per octave from the cutoff points. The resulting sound approximated that of tone pips with a carrier frequency of 2000~.

Fig. 2 shows the changes in sound-pressure produced by the earphones of the pseudophone when the measuring source was placed in any of the five directions indicated (defined with respect to the median plane of the microphones). Comparison of the points of onset of the paired clicks to right and left ears shows the time-differences. Comparison of the amplitudes of their envelopes shows the intensity-differences. The maximal intensity-differences, produced when the source was 80° to the side, were no more than 2 db. Since a 2-db. difference is barely above the limen for a just noticeable deviation of a source from the median plane,[7] binaural intensity-differences were eliminated as cues for directional localization. It

[7] Morgan Upton, Differential sensitivity in sound localization,' *Proc. Nat. Acad. Sci.*, 22, 1936, 409-412.

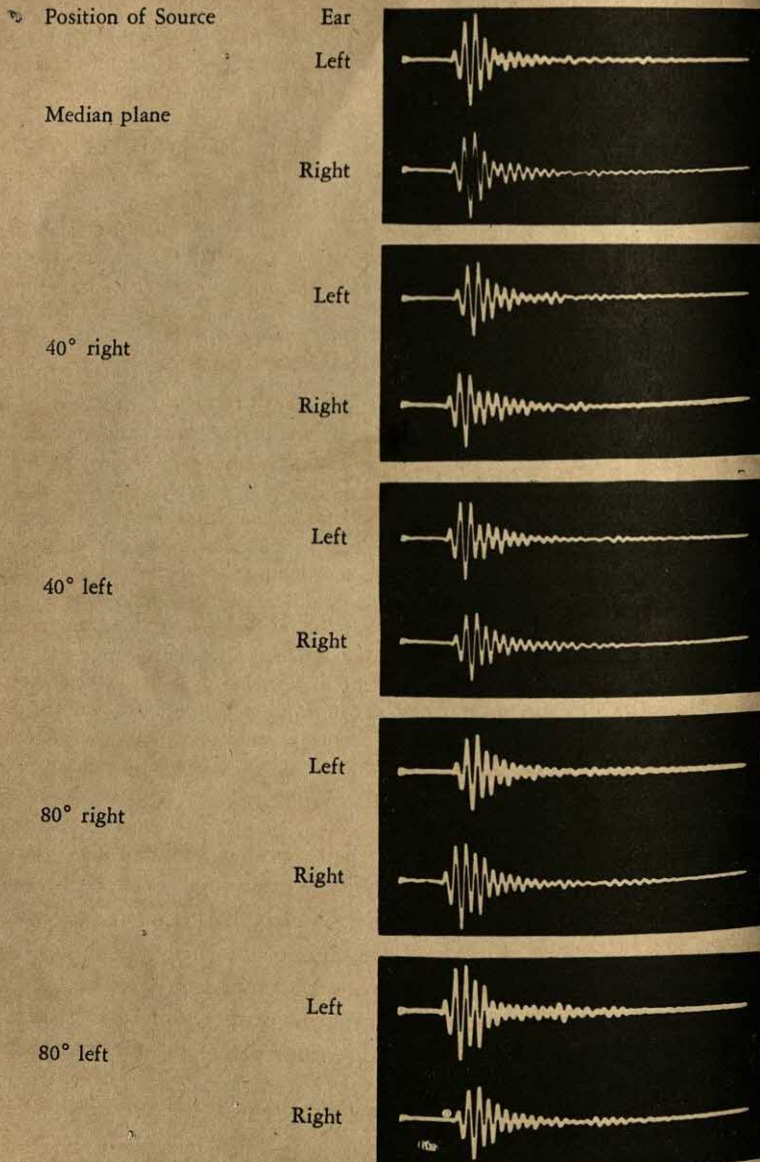


FIG. 2. BINAURAL STIMULI FOR EACH EAR AND POSITION

may be well to point out that although the source of an unmodulated tone of 2000~ is localized poorly, the tone pips as stimuli allow directional discriminations that are almost as accurate as those made to unfiltered clicks. Békésy has shown that the envelopes of such pips determine the time-differences crucial for directional localization.[8] Throughout the experiments the intensity of the measuring stimulus was maintained at 50 db. (sensation-level).

(3) Coupling between earphone and ears. Hollowed insert tips were used to ensure constant coupling both during a session of tests and upon replacements of the pseudophone on *O*'s participation in successive sessions.

As a consequence of the precautions taken for each of the three conditions, shifts of directional localization resulting from differential fatigue of the two ears were avoided,[9] and all considerations of binaural differences could be restricted to the time-differences.

Experiment I. Normal Environment

Procedure. The *O*s were three young men *(A, B,* and *C)* with normal hearing and with experience in psychophysical experiments. *A* and *B* each went through one session lasting about 8 hr.; *C* was used in two such sessions. On the morning of each session, measurements and other tests of localization were taken first with the unaided ears and then with the pseudophone mounted on the head with the microphone-axis normalized. After completion of these tests, the microphone-axis was displaced by 22° and *O*, wearing the pseudophone, was released to go about his normal activities for 7 hr. At the end of this exposure-period, *O* returned to the measuring apparatus, the axis of the microphone was renormalized and another set of measurements was taken under conditions identical to those of the initial set. During the exposures of *A* and *B* and of the first session of C the microphone-axis was so rotated that the left microphone was forward on the head (led the left ear). During exposure in *C*'s second session the right microphone led the right ear. The *O*s remained unaware of the direction and extent of the displacement of the microphone-axis until completion of their sessions.

Results. While walking about in their familiar surroundings, the *O*s noted the qualitative changes in sounds that resulted from transmission through the hearing aids. They did not observe any discrepancies between the apparent auditory directions of sources and their visual positions. When asked, however, to localize a concealed source fixed on the wall of the anechoic room, they made constant errors of directional localization of approximately 22° in the expected direction. When the *O*s moved with

---

[8] Georg von Békésy, Ueber den Knall und die Theorie des Hörens, *Physik. Zsch.*, 34, 1933, 577-582.

[9] Von Békésy, Zur Theories des Hörens: Ueber die Bestimmung des einem reinen Tonempfinden entsprechenden Erregungsgebietes der Basilarmembran vermittelst Ermüdungserscheinungen, *Physik. Zsch.*, 30, 1929, 115-125; Ueber das Richtungshoren bei einer Zeitdifferenz oder Lautstärkenungleichheit der beiderseitigen Schalleinwirkungen, *ibid.*, 31, 1930, 824-835, 857-868.

respect to this stationary source they frequently reported that the sound or its source was moving.

When returned to the measuring device after exposure, all the *O*s began reporting at some time during the measuring either that *there appeared to be two sources or that the source could be localized in two different directions*. The impression of a doubled source was reported when the source was presented within a range of at least 30° in either direction from the median plane of the microphones. After *C*'s second exposure, he designated one apparent source as 'strong' and the other as 'weak.' He proceeded to localize them alternately. The two sets of measurements corresponding to the two directions designated by *C* showed no overlap. The difference in the means was 20.8° when *C* wore the pseudophone, and 18.8° when the measurements were taken with unaided ears. Comparison of these means with the means of comparable preëxposure measurements revealed the following. When *C*, wearing the pseudophone, found the 'strong source' straight ahead of him, the mean direction (from his head) of the measuring source differed by 1.2° to the right from the mean of the preëxposure measurements. When, however, he found the 'weak source' straight ahead of him, the mean direction of the measuring source was shifted 22.0° to the *right* of the preëxposure mean. At this time *C* also consistently localized what he termed a 'middle' or 'centered' direction that proved to lie midway between 'strong' and 'weak' directions.

After *A*'s exposure, he also designated 'weak' and 'strong' sources, and while wearing the pseudophone gave two sets of measurements whose mean differed by 8.5°, a highly significant difference. When he localized what he called the 'strong source' the mean direction of the measuring source differed by 2.2° to the right from the mean of the preëxposure measurements, whereas the measuring source was shifted by 6.3° to the *left* of the preëxposure mean when he localized the 'weak source.'

During the post-exposure measuring of the two other sessions and when *A* localized with unaided ears, the *O*s were unable to localize consistently the 'two sources.' They reported that although the 'weaker source' was apparent in many of the positions they assumed in turning toward the source, it tended to fade out as they approached the position in which they expected it to appear straight ahead. Their estimations of the angle formed between the directions of the 'two' sources were between 10° and 20°. When each *O* was asked to find the best single direction, comparison of preexposure with post-exposure means showed statistically significant shifts that averaged about 10° for localization with the pseudophone and somewhat less with unaided ears. These shifts invariably indicated that the direction

of the measuring source that *O* considered to be straight ahead had become displaced toward the side of the microphone that had led during exposure. Both the shifts and the directions of the 'weak sources' indicated a tendency to correct for the errors of localization caused by rotation of the microphone-axis. Following exposure, *A*, *B*, and *C* in his first session, reported that their judgments were made far more easily and with greater certainty following right turns toward the source than after left turns. Contrariwise, *C* after his second-session exposure reported that his judgments were more easily made after a left turn toward the source.

Discussion. The results outlined above are most easily considered with reference to the idealized schema shown in Fig. 3.

Fig. 3a represents the preëxposure measuring situation. The measuring source was placed alternately to the left SS_L and right SS_R of *O*'s head.

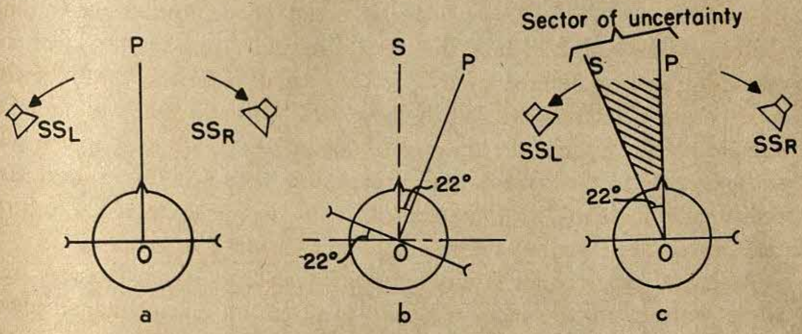


FIG. 3. SCHEMA OF RESULTS
(a) Measuring with axis normal before exposure; (b) Exposed with rotated axis; (c) Measuring after exposure.

O, wearing the pseudophone with axis normal or with unaided ears, turned himself alternately to the left or right around O (the vertical axis) to localize the source. Measurements of these localizations appeared normally distributed about a mean direction P from which a source would produce close to a zero time-difference. The direction P then corresponded to the direction of a source that seemed 'straight ahead' to *O* and was also straight ahead in an anatomical sense. Fig. 3b represents *O*'s head during exposure. In this figure the axis of the microphones, hence the direction *P*, was displaced by 22° clockwise. The direction S then became the anatomically straight-ahead direction. Fig. 3c represents the post-exposure measuring situation in which the axis of the microphones is again normalized. *O*s then found that when the sound-source SS was pre-

sented either in the direction of P, or when displaced by a small angle to the left of P, it appeared to be straight ahead.

Taking into account the various errors involved in measuring and in estimating the effective displacement of the microphone axis, it seems not unlikely that the time-difference produced by sources lying in the direction S during exposure had become a cue for the report of 'source straight ahead' after exposure. In addition, the close-to-zero time-difference from a source in the direction P remained a cue for the report of 'source straight ahead.' Consequently, all *O*s found that there appeared to be two 'straight ahead' directions for one source, and two *O*s were actually able to localize both. The effects of the exposure were not, however, restricted to the interpretation of a novel time-difference as 'source straight ahead.' Rather, the persistent reports that two sources were heard when the measuring source was presented over a wide range of directions implied that a correspondingly wide range of time-differences, including that of zero, had become cues for reports of sources lying in two directions that differed by somewhat less than the expected 22°.

When the *O* represented in Fig. 3c turned to his right to localize SS_R, the direction of the measuring source first coincided with P. When, however, he rotated to his left to localize SS_L, the direction of the source first coincided with S, the direction of the source that produced the cue for the newly established response. In the latter case the *O*s reported far greater uncertainty. Moreover, the *O*s who could not separately localize the two apparent sources after exposure tended to compromise on a direction lying within the sector bounded by S and P.

Similar effects were produced in preliminary experiments in which *O*s were exposed with a larger rotational displacement (45°) of the axis of the microphones. The post-exposure measurements were, however, considerably more erratic—presumably a result of *O*'s difficulty in making compromise judgments. This consideration may partially explain the failure of the investigators who used the radical 180° displacement of the aural axis to obtain any evidence of shifts of localization under conditions comparable to those of this experiment.[10]

Experiment II. Controlled Conditions

The normal environment provides a large number of conditions of which any one or any number may be responsible for changes in directional

[10] P. T. Young, Auditory localization with acoustical transposition of the ears, *J. Exper. Psychol.*, 11, 1928, 399-429; C. F. Willey, Edward Inglis, and C. H. Pearce, Reversal of auditory localization, *ibid.*, 20, 1937, 114-130.

localization. To find the critical factors, analysis and control of the conditions of exposure were necessary. *O*s with a 22° rotation of the axis of the microphones did not report discrepancies between the visual and auditory directions of familiar sound-sources even upon close questioning. Their lack of such discrimination suggested that *O*'s awareness of a discrepancy was not a necessary condition for producing the shifts. This conclusion was substantiated by the shifts of localization produced in preliminary experiments in which blindfolded *O*s walked about in the presence of sets of sound sources rigged so as to preclude completely any recognition of their errors of localization. It appeared then that although *O*'s shifts of localization tended to correct for the induced errors, these shifts were not dependent upon *O*'s information about the errors. If this conclusion was correct, the shifts and the apparent doubling of the source could have occurred only as a consequence of *O*'s exposure to combinations of stimuli that differed from those he ordinarily experienced. This difference, resulting from the displacement of the aural axis, consisted of a systematically changed relation between acoustic and movement-produced proprioceptive stimuli.

A number of investigators have shown that certain kinds of proprioceptive excitation given simultaneously with acoustic stimuli may influence the localization of sound-sources.[11] From the experiences of aviators exposed to unusual inertial forces, we have learned that variations in proprioceptive (labyrinthine) stimulation can systematically change the directional responses of a listener even though the acoustic stimuli remain constant. Quantitative information on these phenomena has come from experiments designed to produce similar effects by the use of a human centrifuge.[12] Wallach,[13] whose work has been partially confirmed by Klensch,[14] demonstrated that the covariation of acoustic and proprioceptive stimuli, involved in rotational displacements of a listener's head in the presence of a sound-source, serves as a cue for accurate localization of the source in three dimensions. These findings imply that the cues adequate for localization are always certain combinations of proprioceptive and acoustic stimuli.

From geometrical and physical considerations it is obvious that the time-difference, which functions as the binaural cue *par excellence,* cannot itself give sufficient information to determine unambiguously the location or even direction of the source that produces it. From the same considera-

---

[11] For a short review of some of this work see M. D. Arnoult, Post-rotatory localization of sound, this JOURNAL, 63, 1950, 229-236.

[12] Brant Clark and Ashton Graybiel, The effect of angular acceleration on sound localization: the audiogyral illusion, *J. Psychol.,* 28, 1949, 235-244; Ashton Graybiel and J. I. Niven, The effect of a change in direction of resultant force on sound localization: the audiogravic illusion, *J. Exper. Psychol.,* 42, 1951, 227-230.

[13] Wallach, *op. cit.,* 339-368.

[14] Klensch, *op. cit.,* 492-500, 706-713.

tions, however, knowledge concerning head-displacement in the presence of a sound-source and the accompanying sequence of changing time-differences can provide sufficient information. In a space containing a stationary sound-source, a given displacement of the head and ears from an initial position causes a physically determined sequence of time-differences. With trivial exceptions the combined information (displacement plus sequence of time differences) uniquely determines the position of the source relative to the head. Hence to the extent that excitation of the proprioceptors corresponds to displacements of the head, the combinations of stimuli can in principle provide information for unambiguous localization by a listener.

The orientation of the aural axis on the head may be considered a parameter—normally constant—among the geometrical variables that determine the combinations of stimuli which may occur. After rotational displacement of this axis novel combinations of stimuli become the rule and localization becomes erroneous. A simple illustration of a novel combination of stimuli is shown in Fig. 4. In Fig. 4a, the head of *O* with normal aural axis is shown moving from A to B. Under this condition the initial time-difference at A is zero, and the difference remains at zero throughout a translational displacement to B. In Fig. 4b, the aural axis of *O* is displaced. An initial time-difference of zero is produced by the source, but the same translation from *A* to *B* causes a continuous change of the time-difference. It was under the latter condition that the *O*s of Experiment I reported movement of a source normally perceived as stationary.

Under natural conditions any attempt to specify exhaustively these combinations of stimuli is greatly complicated by the existence of moving sources of sound. Nevertheless, it seems reasonable to distinguish between *typical* and *atypical* combinations of stimuli.[15] *Typical* describes combinations of stimuli that accompany the motion of a listener in a space containing sound-sources. *Atypical* describes combinations of stimuli that accompany the motion of a listener whose aural axis has suffered displacement. Prolonged exposure to atypical combinations of stimuli appears to be a necessary condition for producing the shifts of localization.

In Experiment I no attempt could be made to specify the incidence of

---

[15] The use of these terms is meant to convey the idea that the two classes of combinations are by no means mutually exclusive under the most general conditions of exposure. The distinction implied is a statistical one. An explication of the meaning of these terms must ultimately be made in terms of a mathematical or statistical model, derived from reasonable assumptions about the distributions, in a normal environment, of the variables that determine the combinations of stimuli.

atypical stimuli that occurred during exposure. A preliminary finding that significant shifts could be produced in 1 hr. or even less made feasible a factorially designed experiment. In the following experiment the combinations of stimuli to which *O*s were exposed were controlled in an attempt

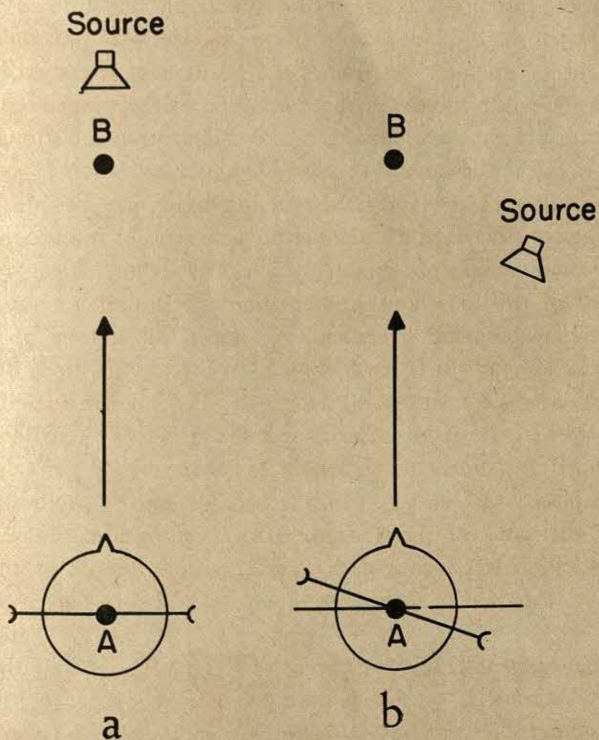


FIG. 4. ILLUSTRATION OF NOVEL COMBINATION OF STIMULI
In (a) head translated from A to B, time-difference remains constant. In (b) head with rotated aural axis translated from A to B, time-difference varies.

to discover some of the aspects of atypical stimulation that are essential for producing the shifts.

Procedure. A degree of control over head movements was achieved by having *O* walk a fixed path during the entire period of exposure. The experiment was carried out in the anechoic room shown schematically in Fig. 5. At each of the positions A, B, and C, ¼-w. lamps were placed. By means of a timing circuit, one lamp at a time was turned on for 5 sec., after which the next lamp in the sequence went on for 5 sec. The lamps were turned on in the repeating sequence ABC, and *O*, who was instructed to walk toward the lamp that was lit, walked

continuously along the triangular path. To assist him in keeping a constant rate of walking, a break was introduced into the sound once per sec. Keeping in step with this rhythm, O walked five steps from A to B, made a left turn of 120°, walked another 5 steps to C, turned, and so on. With the sequence ACB, O followed the same path in the reverse direction, making a right turn at each corner. These lamps and a well-shaded reading lamp used by E were the only sources of illumination in the room during exposure.

At each of the A, B, and C positions, a sound-source (Permoflux PDR-8 earphone) was attached to the wall at a height of 6 ft. These sources were actuated by the

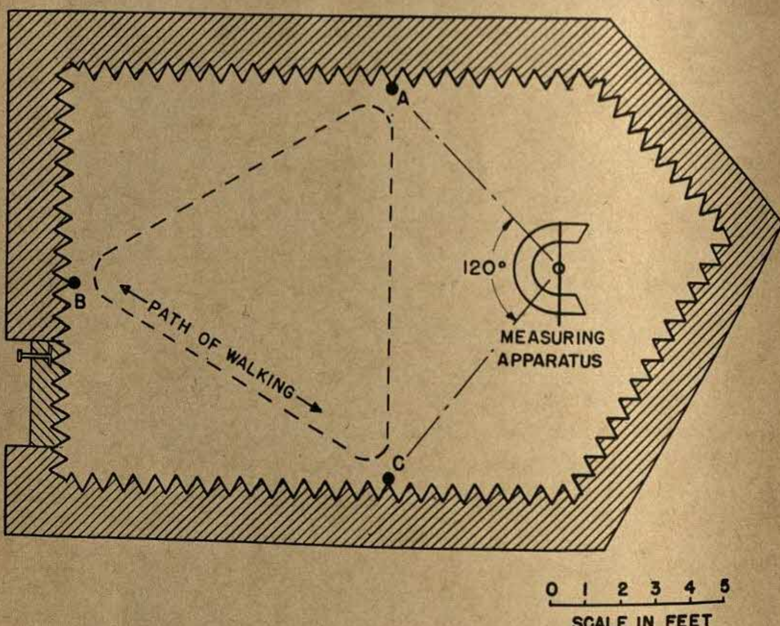


FIG. 5. PLAN OF THE ANECHOIC ROOM
A, B, C, positions of sound-source and lamp.

filtered clicks previously described and followed the same temporal on-off sequence as the lamps. The relative phasing of the lamp and sound circuits could, however, be adjusted to produce any one of three combinations. The lamp in any one corner might: (1) lag behind its adjacent source by 5 sec.; (2) be turned on simultaneously; or (3) lead by 5 sec. By this means O could be made to walk each leg of the triangle while the sounding source was (1) always *lateral* to his line of translation, called Condition L (as when O walked from A to C while Source B was on); (2) always *ahead* and in his line of translation, called Condition K (as when O *walked* from A to C while C was on); or (3) always *behind* and in his line of translation, called Condition M (as when O walked from A to C while A was on).

The concurrent pairs (source and lamp) were actuated simultaneously. Con-

sequently, O always commenced the 120° turn at each corner of the triangle in time to approach the next-lighted lamp while its concurrent source was sounding. This procedure ensured that the position of the source was consistent during O's major head rotations. Small oscillations of the head occurred continually while O walked. The three variations of the source-position with respect to O's line of translation (L, K, and M) together with the two alternative directions of turn (sequence ABC and sequence ACB) resulted in six alternative combinations of translational and rotational movements of the head with respect to a fixed source. In addition to these, the effects of rotation alone (called *no translation* or NT) could be assessed

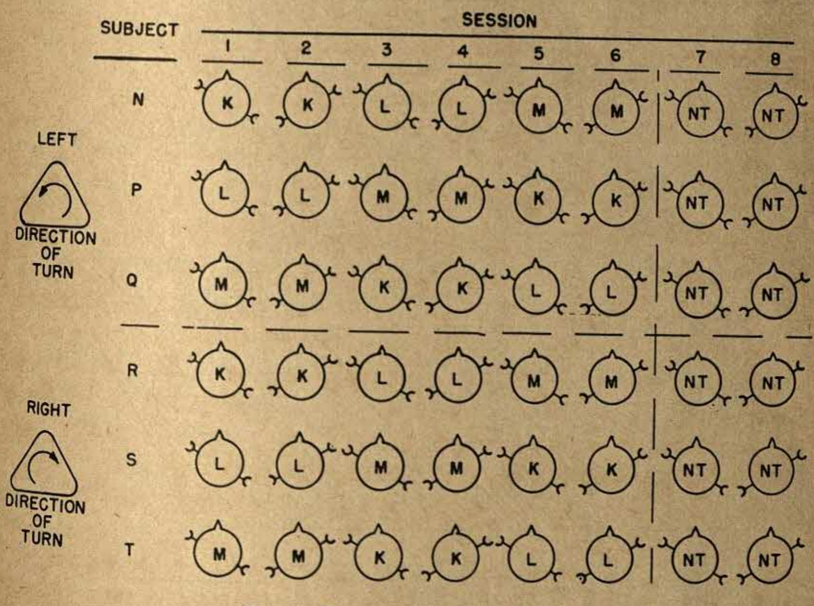


FIG. 6. PLAN OF EXPERIMENT II

K, source ahead in line of translation; L, source lateral to line of translation; M. source behind in line of translation; NT, no translation.

by having O, seated in the rotating chair as in the measuring procedure, rotate back and forth through 120° every 5 sec. in response to lamps placed at A and C. In this case each source and its adjacent lamp went on and off simultaneously.

The design of the experiment is shown schematically in Fig. 6. The Os were six young men, all with normal hearing and with ears matched in sensitivity within 10 db. None knew the purpose of the experiment and they received no information concerning the displacement of the microphone-axis. After preliminary training each O went through eight sessions, separated from each other by several days. Three of the Os (N, P, and Q) walked in the sequence ABC the other three (R, S, and T) walked in the sequence ACB. Each of the K, L, and M conditions was given in successive pairs of sessions. The only difference between members of each

pair was displacement of the axis of the microphones in the opposite sense. The temporal order of the sessions K, L, and M was permuted among each group of three *Os* to balance practice-effects. The NT sessions were always the final two of the eight sessions.

During each session the total exposure-period of 1 hr. was divided into three periods of 20 min. each, interspersed with measurements of directional localization. Thus, four sets of measurements were taken during each session; one initially, and one after successive periods of 20, 40, and 60 min. Each measuring set consisted of 14 successive localizations taken while *O* wore the pseudophone with the axis of the microphone normalized. *O* was instructed to align himself quickly to the position in which the source sounded straight ahead of him. Before the initial set of measurements was taken, *O* was fatigued by exposure to the measuring stimulus at zero time-difference for 4 min. This procedure was intended both to prevent qualitative differences between the sounds heard before and after exposure, and to eliminate the influence of differential fatigue resulting from exposure to any residual binaural intensity differences. The period of 4 min. was chosen on the basis of Hood's report that the effects of such differential fatigue (he calls it perstimulatory) upon binaural comparisons of loudness reach an asymptote in approximately 4 min.[16]

Results. Each session, except 7 and 8, involved approximately equal amounts of translational movement of *O*'s head. The consistency of the results obtained among the *Os* justifies their comparison along with comparisons of results obtained in the several sessions of a single *O*. The results are plotted on the graphs of Figs. 7, 8, and 9. The directional shift was the angular difference between the mean preëxposure direction of the measuring source from *O*'s head and the means of the post-exposure directions. The shift was considered positive (or negative) when *O* turned his head farther to the left (or right) to find the measuring source straight ahead of him after exposure. Under the conditions of measurement, a positive (or negative) shift corrected for the error of localization imposed by the pseudophone during exposures in which the right (or left) microphone had been forward on the head. The standard deviations of the means (plotted points) of sets of measurements varied between 2° and 5°. The legend of each graph indicates for each point the *O* and session number found in Fig. 6.

Exposure with lateral source (L). Figs. 7a and 7b each show the results of six sessions, one from each of 6 *Os*. The exposure-conditions whose results are plotted in Fig. 7a were distinguished from those of Fig. 7b only by the sense of rotation of the microphone-axis. Because the left and right turns did not appear to influence the results, as judged by the inter-

---

[16] J. D. Hood, Studies in auditory fatigue and adaptation, *Acta Oto-laryng.*, Suppl., 92, 1950, 1-56.

mingling of filled and unfilled points, the data taken under these two conditions have been pooled. The mean shifts show neither an appreciable shift of directional localization nor a significant difference caused by having left as opposed to right microphone forward on the head.

Exposure with rotation alone (NT). Figs. 8a and 8b each show the results of six sessions, one from each of 6 *O*s. The exposure-conditions

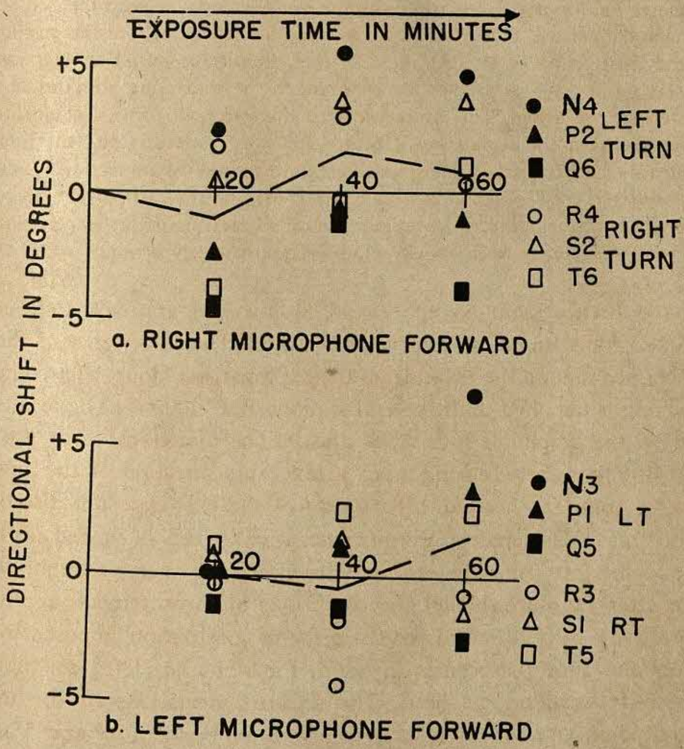


FIG. 7. EXPOSURE WITH LATERAL SOUND-SOURCE

whose results are plotted in Fig. 8a were distinguished from those of Fig. 8b only by the sense of rotation of the microphone-axis. The mean shifts average about 1° in directions that tend to correct for the error of directional localization imposed by the rotation of the microphone axis.

Exposure with source ahead and source behind (K and M). Figs. 9a, 9b, 9c, and 9d each show the results of six sessions, two from each of 3 *O*s who were exposed to both a sound-source ahead (filled points) and a sound-source behind (unfilled points). The data taken under these two

conditions are combined in single graphs, irrespective of the sense of rotation of the microphone-axis during exposure. Of the conditions combined: (1) right microphone forward—source ahead with left microphone forward—source behind, and (2) left microphone forward—source ahead with right microphone forward—source behind; each involved an un-

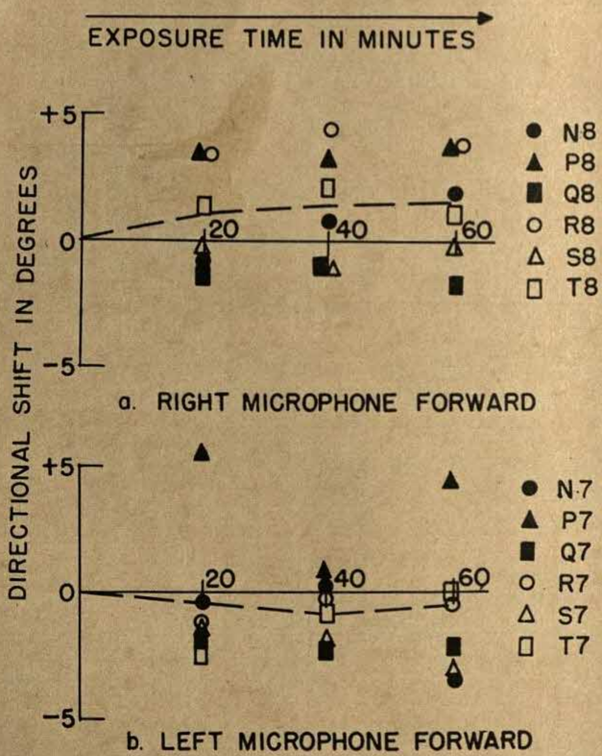


FIG. 8. EXPOSURE WITH ROTATION ALONE

derlying factor that proved decisive for the direction of shift. This common factor was the ear in which the sound had led—each separate click arrived first—during exposure. The intermingling of filled and unfilled points especially evident in Figs. 9a and 9d indicated that the same shifts were produced despite the apparent difference in conditions of exposure. The results in Figs. 9a and 9d indicate shifts of close to 10°. The shift appears to increase rapidly at first and to reach a maximum at the end of an hour.

In addition to the quantitative results, the reports of the *O*s obtained under the conditions that yielded the largest shifts tended to agree with

those reported in Experiment I. Several *O*s reported that they made their judgments more easily when they turned to localize the source toward the side of the ear in which the sound had lagged during exposure. Almost

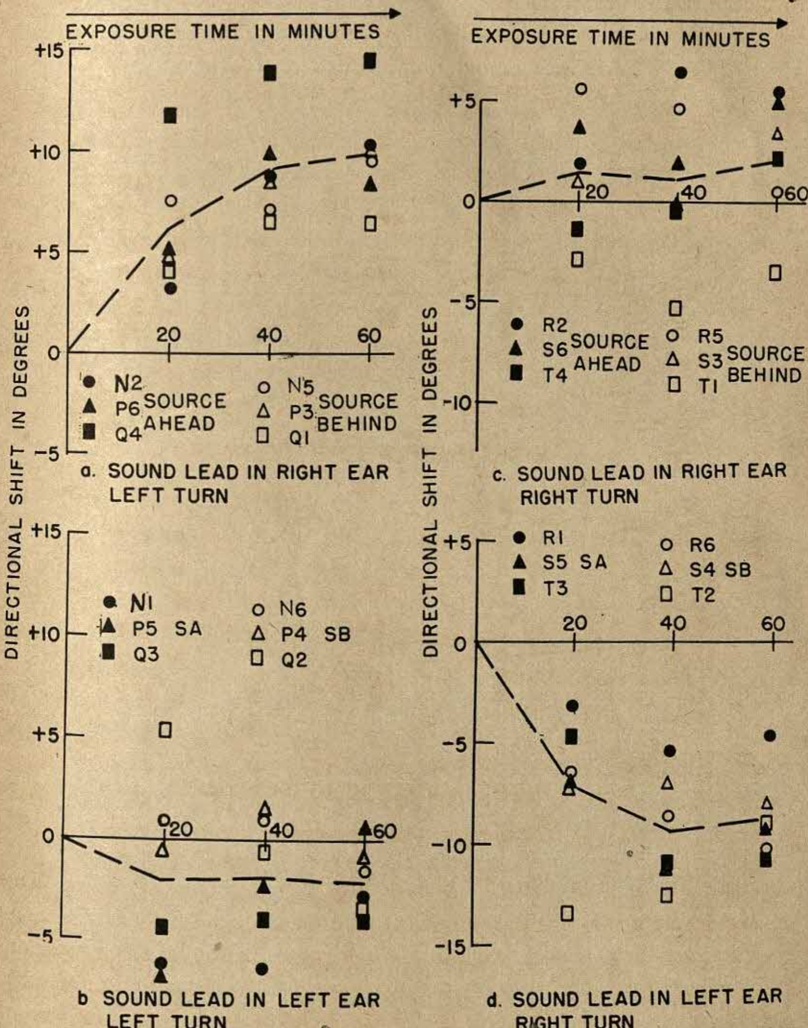


FIG. 9. EXPOSURE WITH SOUND-SOURCE AHEAD AND BEHIND

all the *O*s reported that the source appeared to be more spread out in the horizontal plane after exposure. During a post-exposure measuring period one *O* asked if there were not two sources present.

Discussion. Two of the conditions of exposure that produced the data (filled and unfilled points) combined in Figs. 9a and 9c are depicted schematically in Figs. 10a and 10b. When the microphone-axis was so displaced that the right microphone was forward, *O* walked toward a source that was invariably ahead of him. When, however, the left microphone was forward, the source was behind him. Inspection of these two situations

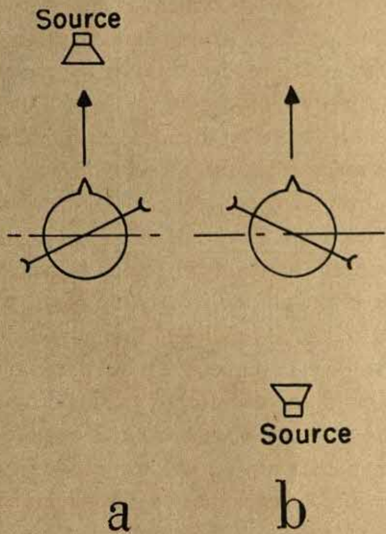


FIG. 10. COMBINED CONDITIONS
(a) right microphone forward with sound-source ahead; and (b) left microphone forward with sound-source behind are identical with respect to the ear in which the sound led during translation.

shows that they are identical with respect to the ear in which the sound led during translation. A similar identity held for two of the conditions of exposure that produced the data combined in Figs. 9b and 9d with the exception that the sound led in the left ear. Despite the fact that the sound-sources were reversed in position from front to back, when in combination with opposite rotational displacements of the axis of the microphones they caused the same binaural differences to accompany the translational movement depicted.

Consideration of the conditions depicted in Figs. 10a and 10b in terms of the direction of shift required for correction of the imposed error shows that the former case required a positive shift, the latter, a negative shift. Since the results show that in both cases a positive shift occurred, the condition with sound-source behind has actually resulted in anti-corrective

shifts. Under these special experimental conditions, the imposed error of directional localization was *increased* by exposure. This finding lends further weight to the previously expressed view that the shifts do not represent some kind of purposive correction but rather are strictly a consequence of the combinations of stimuli to which *O*s are exposed in the course of their movements.

Comparisons of the results represented in Fig. 9a with those of 9c, and of the results in 9b with those of 9d show that large differences in the magnitudes of the shift were produced under conditions that differed only in the sense (left or right) of the major head rotations. In both comparisons, the larger shifts were produced when the sense of these rotations was in the direction away from the ear in which the sound led.

The shifts of localization result from directions of response that appear to be compromises between the preëxposure, normally accurate localizations, and newly acquired directions of response. Increase of the shift with duration of exposure must indicate an increase of the strength of the newly acquired direction of response relative to the normal direction of response. After long exposures—as in Experiment I—the two response directions tend to become separated and *O* interprets a single time-difference as a cue for two possible directions of a sound-source. Some of the characteristics of atypical stimulation that are critical for producing the shifts may be found by analysis of the combinations of movements of *O*'s head with the time-differences that accompanied them under the conditions of this experiment.

If the changes of the time-difference caused by head rotations are considered partialed out, then under the K and M conditions the time-differences were relatively invariant with translation of the head. Under the L condition, however, the direction of the sound-source from *O*'s head continually varied with translation and consequently the time-differences varied even if the effects of head rotation are considered partialed out. When *O* walked one leg of the triangle he continually made small rotations of his head over a range of approximately 30° to the right and left while the center of his head translated in a relatively straight line by the saccadic displacements characteristic of walking. The relevant proprioceptors are probably responsive only to motion defined with respect to the coördinates of the head and body at the initiation of a movement or acceleration. Hence under the K and M conditions each of a set of translational displacements was invariably combined with a unique time-difference and these combinations continually recurred. Under the L condition this invariance was lacking. The results obtained after exposure under the NT condition

indicate that the mere prolonged recurrence of fairly constant time-differences in the absence of translation of the head is insufficient to produce the large shifts. Hence, the invariant combination of time-difference with translational displacement appears to be the factor in atypical stimulation that is critical for producing the shifts of localization and related effects. In addition, certain combinations of rotational movement and accompanying time-differences play an important secondary rôle which cannot be specified without further experimentation.

Conclusions

Let us assume that any directional localization is a surrogate or equivalent response for a translational displacement of the head in the same direction. In accord with the above discussion we may then say that continued recurrence of the combination of a certain direction of head-translation with a given time-difference will eventually result in the recapitulation of that translational displacement when *O* is asked to localize a sound-source producing the given time-dfference at his ears. As a demonstration of this rule, the experimental results are complicated by the presence of the normal preëxposure directional responses. In most cases these appear to cause compromise directions of response evidenced as small shifts of localization. The results indicate, however, that prolonged exposure to the atypical stimulation finally builds up the strength of the new response to the point where doubled localizations are obtained. Hence, the above rule, when qualified by a statement about the strength of the new response relative to the strengths of other more-or-less probable responses, appears to be adequate to account for the experimental results. It may also, however, have relevance for the more general problem of the original acquisition of localization.

Under the natural conditions of exposure (Experiment I) the shifts and doubled localizations tended to restore the accurate localization that was made erroneous by rotation of the aural axis. This result, in light of the above discussion, implies that such exposure entailed the critical combinations of stimuli in a manner that tended to result in correct localizing responses. Under these atypical conditions of exposure, each time-difference was presumably combined more frequently with a translation of the head in the direction of the source producing this time-difference than with any other direction of translation. By similar reasoning, if it could be shown that the neonate in the course of his movements is exposed to any selected time-difference combined most frequently with a particular direc-

tion of head translation, the directional response to the selected time-difference could be predicted. If each time-difference or, more realistically, each of some subliminal range of time-differences could be shown to have been most frequently combined with a direction of translation identical with the direction of the source producing this time-difference, a theory of acquisition by exposure would be plausible. It is obvious that as a direction of translation of the head approaches the direction of a source, the time-difference varies less for a given displacement of the head. An adequate formulation of this notion might predict the acquisition of normal directional localization. The task involves essentially a mathematical description of typical combinations of stimuli and a comparison of their probabilities of recurrence under normal environmental conditions. A preliminary formulation, developed elsewhere,[17] shows that some of the properties of normal localization may be derived from these relatively simple considerations.

SUMMARY

Children maintain accurate sound-localization despite growth of the head which systematically changes the binaural cues produced by sound-sources in constant directions from the head. If, as suspected, this phenomenon is conditioned by exposure to the environment, then similar systematic changes suffered by an adult, may produce shifts of localization during exposure. In the experiments an electronic pseudophone, worn by adult *Os*, displaced the aural axis by 22° around the vertical axis of the head. The effects were studied by comparing measurements of direction finding taken before, during, and after exposure under this condition of hearing.

In Experiment I, 3 *Os* with displaced aural axes spent periods of 7 hr. in their normal environments. At the end of this period when a single source was sounded each *O* reported either that there appeared to be *two* sources or that the source could be localized in *two* different directions. Comparison of measurements showed shifts of localization averaging about 10° and separations of the two apparent sources ranging from 8° to 21°. Although the *Os* remained unaware of errors of localization caused by the rotation of their aural axes, the shifts tended to correct for these errors.

To account for the results a perliminary analysis was made of the changed relation between acoustic and proprioceptive stimuli caused by

---

[17] Richard Held, Trained shifts in binaural direction-finding: their implications for the genesis of auditory space. Unpublished doctoral dissertation, Harvard University, 1952.

displacement of the aural axis. The notion of *typical* as opposed to *atypical* combinations of stimuli was introduced to characterize the changed conditions.

In Experiment II the conditions of exposure (motion of *O* and location of sound-sources) were controlled in an attempt to discover some of the aspects of atypical stimulation that produce the shifts. Each of 6 *O*s went through 8 1-hr. sessions of controlled exposure. Fourteen conditions of exposure were presented in a factorial design. The results show that invariance of the time-difference with translation of the head in a particular direction is the critical factor under the experimental conditions. Shifts averaging 10° were produced in 1 hr. under the optimal conditions of exposure.

The generalizations advanced to deal with the results of these experiments may be relevant to the more general problem of the original acquisition of auditory localization.

NEUROMUSCULAR CONTROLS IN MAN: METHODS OF SELF DIRECTION IN HEALTH AND IN DISEASE

By EDMUND JACOBSON, Laboratory for Clinical Physiology, Chicago, Illinois

Sooner or later the responsibility of teaching people how to conserve their energies in the interests of health and of personal efficiency naturally falls upon psychologists as well as upon physicians. The approach can be in the name of common sense, advising the individual on rest and vacations, without really showing him how to rest; or it may take on a more scientific character in the form of technical instructions on (1) how to relax physiologically, generally and differentially, or (2) how to run his organism in daily living. Procedures relating to the first-mentioned have been described previously and evidence of the interrelationship of neuromuscular and of mental activities has been fully discussed.[1] It has been shown that with progressive muscular relaxation (a) the knee-jerk and other reflexes dwindle or disappear[2] along with (b) mental activities, including emotions, having generalized neurogenic deactivation effects on (c) the vegetative nervous system, in which the cardiovascular,[3] the gastrointestinal and other systems participate.[4]

Instruction by the method of control, the second form mentioned above, rests upon evidence secured during the last four years from more than 110 patients and healthy subjects. The new approach differs from the old and appears to be more readily and effectively learned. The method of *general* relaxation shows the individual how to slow his energy expenditures and

---

* Accepted for publication October 13, 1954.

[1] Edmund Jacobson, Progressive relaxation, this JOURNAL, 36, 1925, 73-87; *Progressive Relaxation,* 1938, 164-189, 327-345; Electrical measurements of neuromuscular states during mental activities, *Amer. J. Physiol.*, 91, 1930, 567-608; 94, 1930, 22-34; 95, 1930, 694-702; 95, 1930, 703-712; 96, 1931, 115-121; 96, 1931, 122-125; 97, 1931, 200-209; Electrophysiology of mental activities, this JOURNAL, 44, 1932, 677-694; Electrical measurements of mental activities in man, *Trans. N.Y. Acad. Sci.*, 2:8:8, 1946, 272-273.

[2] Edmund Jacobson and A. J. Carlson, The influence of relaxation upon knee-jerk, *Amer. J. Physiol.*, 73, 1925, 324-328.

[3] Jacobson, Variation of blood pressure with skeletal muscle tension and relaxation, *Ann. Int. Med.*, 12, 1939, 1194-1212; The heart beat, *ibid.*, 13, 1940, 1619-1625; Variation in blood pressure with skeletal muscle tension, *Amer. J. Physiol.*, 1939, 546-547; Variation of blood pressure with brief voluntary muscular contractions, *J. Lab. & Clin. Med.*, 25, 1940, 1029-1037; Cultivated relaxation in 'essential' hypertension, *Arch. Phys. Ther.*, 1940, 21, 645-654.

[4] Jacobson, Spastic esophagus and mucous colitis, *Arch. Int. Med.*, 39, 1927, 433-445.

is chiefly, although not exclusively, on the negative side. The positive side is obvious in *differential* relaxation[5] but is emphasized in the method of control. It not only envelops the method of relaxation, but constitutes a new approach. Instructions are based on the following considerations.

A new point of view. In the sciences of man, a distinguishing feature can often be discerned from a new angle. The feature with which we are concerned is the possibility of man acquiring greater ability to run his organism in its best interests. Just as a valley may be surveyed from the top of different hills, so we may look upon the activities of man not solely from the conventional points of stimulus and of response, of instinct and impulse, of reflex and habit, but also from the point of view of his opportunity to run his organism, which he never before had learned to do.

We are familiar with various man-made machines and instruments possessing certain devices for the government or control of the operation of the mechanism as a whole. The automobile is an example. While this vehicle moves under its own power, it is not self-directing, but the motion of its wheels can be determined in rate and in direction by devices operated by the driver. These devices are often called the 'controls' and their presence suggests that we look for analogs in the human mechanism.

Corresponding to the wheels of the car are the muscle tissues in man. Just as motion of the car depends directly upon the revolution of the wheels, so action of the human organism as a unit depends as a rule directly upon the shortening of muscle fibers while relaxation includes their lengthening. By these means alone man can move and can effect divers changes in his environment and in himself, in accordance with the first law of Newton, which applies likewise to the motions of the automobile.

The ratio of the work necessary to operate controls (C) of an automobile to the work performed by the wheels (E) can be represented by the fraction C/E. Automobile engineers generally have directed their efforts towards making this fraction smaller with each year's model, partly by increasing the value of E but chiefly by lessening the value of C. Assuming that evolution tends in the same direction, we should expect in man an extremely small value for C. In consequence of the slight energies involved, the very existence of C might tend to be overlooked.

If we represent the varying stages of evolutionary development attained by organisms of different species and genera by corresponding fractions C/E, a rough conception is secured of the work-gain attained in each instance. In one, notably the mastodon, the value of E evidently became extraordinarily large. Evolutionary history indicates the failure of such organisms to survive, based upon attainment of excessive values of this denominator. The greater success attained in the evolution of man can be represented more clearly if we invert the fraction, making it E/C, which roughly indicates the efficiency of the mechanism of control. Just as automobile engineers found it best to limit the power delivered at the wheels, so evolutionary tendencies evidently have met with the most marked success in man, whose effector strength is far less than that which characterized the mastodon. In man the ratio

[5] E. J. Gunderson, Relaxation in therapy, *J. Gen. Psychol.*, 38, 1948, 181-190.

E/C has become relatively large, but most striking has been the *qualitative* change in the manifestations of *E*. Man is able to carry out many and varied types of activities in addition to the forward and backward locomotion characteristic of the automobile. Among these are abilities to see and to imagine seeing, to hear and to imagine hearing, to write and to imagine writing, and to recall correspondingly. Evidently in the course of evolution, *E* has become diversified into many functions accomplished through the shortening and lengthening of muscle fibers which rotation of wheels could not effect. As will be illustrated subsequently, the diversified abilities of man derive from the favorable location of *C* all over the exterior surface of his body beneath his skin, rather than exclusively in one locality as in the instance of the automobile, where they are located on the dashboard. Thus via the skeletal musculature he can direct activities of his various parts due to devices which (1) indicate to him the nature and the magnitude of energy expenditure at any portion of his external organism and (2) enable him to initiate and to direct the movement of his individual parts in a way that is obviously impossible for any instrument or machine constructed by man up to date.

When a switch or other mechanical device is operated by man, the control device is separated in space from the operator. A different set of conditions obtain if, when, and as man operates autogenous controls, for here the control device (motor nerves, muscle spindles) and the control indicators (muscle sensations) are not separable from the operator himself, but both operator and control devices and control indicators subsist in one and the same organism. To this extent man's employment of autogenous controls departs from strict analogy with his use of dashboard controls, for instance, in driving an automobile. To the extent that autogenous controls are inseparable from the operator it is inaccurate to draw a strict analogy between man's possible operation of himself with that of an automobile and the deviation from accuracy could be regarded as error of a constant variety. Examples of departure from strict accuracy abound, however, among scientific conceptions which nevertheless often prove not only of practical utility but also of value for scientific understanding providing that any error or deviation is duly noted and properly allowed for (Fig. 1).

Man's control of his organism. We must revise the traditional concept that the machinery of human ideation lies exclusively in the brain and other parts of the nervous system. It is necessary to consider that in no moment of ideation, emotion, or other human activity does any nerve shorten, or lengthen, or engage in any structural change of location whatsoever. The brain and other nervous tissue are capable of conducting electrochemical impulses only, and in themselves are incompetent to effect direct mechanical change internally or on the environment. It has been shown from this laboratory that mental operations, including imagination, recall and all forms of emotion, depend upon specific operations of the skeletal (and other) musculature which doubtless are approximately simultaneous with (rather than subsequent to) hypothesized specific brain action (Fig. 2). While locomotion of the organism or of its parts involves mus-

cular contractions which are readily recognized because they can be seen with the naked eye, mental operations involve muscular contractions not readily recognized because they are microscopic. They thus require mechanical or electrical amplification for identification. Employing methods of measuring action-potential previously developed[6] (Fig. 3), we now add evidence which indicates that within certain limits man is provided with

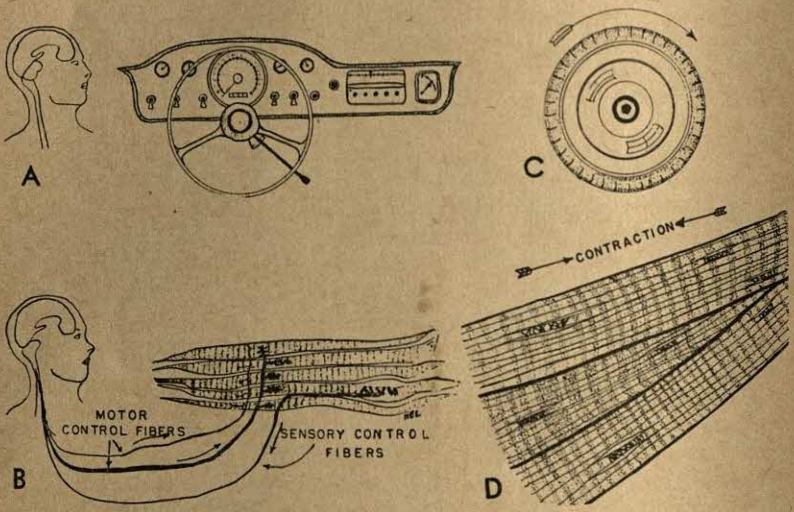


FIG 1. COMPARISON OF MECHANICAL AND AUTOGENOUS CONTROLS

The driver (A) makes contact with external control devices. By employing these mechanical controls he can determine rate, direction, and duration of car motion. The automobile's wheels are the effector mechanisms subject to control devices in or near the dashboard. The patient (B) is in constant contact with internal control devices, which he can be trained to observe and to employ purposively. By employing these physiological controls he can determine degree, location, extent, and duration of the activity of the organism as a whole. Man's shortening muscle fibers are the effector mechanisms subject to control devices in or near the muscles themselves. These control devices include large and small motor nerve fibers; also sensory fibers from muscle spindles.

psychophysiological functions such that through certain technical training he can acquire something like a driver's control of his organism and of

[6] Jacobson, The neurovoltmeter, this JOURNAL, 52, 1939, 620-624; The direct measurement of nervous and muscular states with the integrating neurovoltmeter, *Amer. J. Psychiat.*, 97, 1940, 513-523; An integrating voltmeter for the study of nerve and muscle potentials, *Rev. Sci. Instruments*, 11, 1940, 415-418; Recording action-potentials without photography, this JOURNAL, 54, 1941, 266-269; Edmund Jacobson and F. L. Kraft, Contraction potentials in man during reading, *Amer. J. Physiol.*, 137:1, 1942, 1-5; The cultivation of physiological relaxation, *Ann. Int. Med.*, 19, 1943, 965-972.

various portions (Fig. 4). Such control derives from the muscle sense which functions with skeletal muscular contraction and with certain 'servo' mechanisms, chiefly of visual character. Progressive relaxation—the negative phase of human self-direction—has been discussed. Previous communications support the view that the self-direction of man can include his mental operations insofar as these are determined by minute

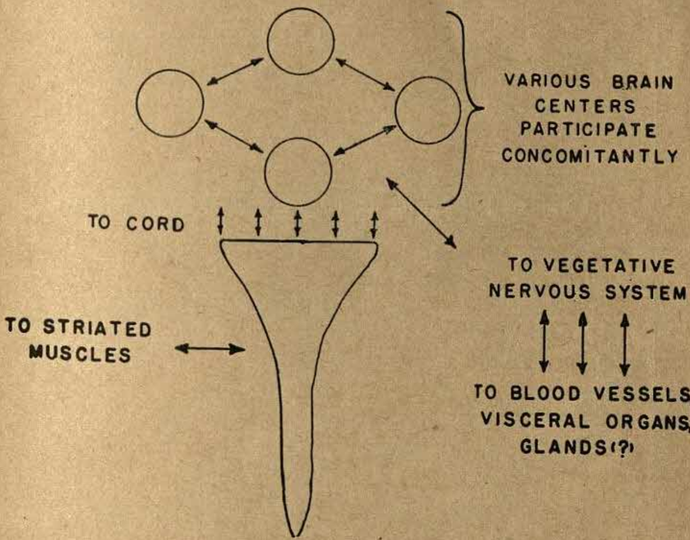


FIG. 2. DIAGRAM OF SYSTEM ACTIVATION

During (1) overt behavior and during (2) mental activity in man, present day evidence supplants the traditional view that the brain acts first (idea) and then the body or vice versa (William James re: emotion). Activation in the total nervous system, brain included, occurs with two-way discharges *approximately simultaneous* with activation in the peripheral organs innervated. Arrows indicate the approximate temporal concomitance.

contractions of the eyes, of speech, or of other parts specifically involved in each mental act.

In no sense does this imply that the nervous system is unnecessary for reflection or even for locomotion. If we needed any evidence against this view, which we do not, the incapacity of children born anencephalic would be conclusive. The present thesis is only that in certain views which have become traditional, functions which only muscle tissues could carry out have been vaguely attributed to nervous tissue, thus hindering the progress of the sciences relating to the nature of human action.

Consequences. To clarify what has been said above, we can ask just what may healthy and diseased individuals derive from the proposed tech-

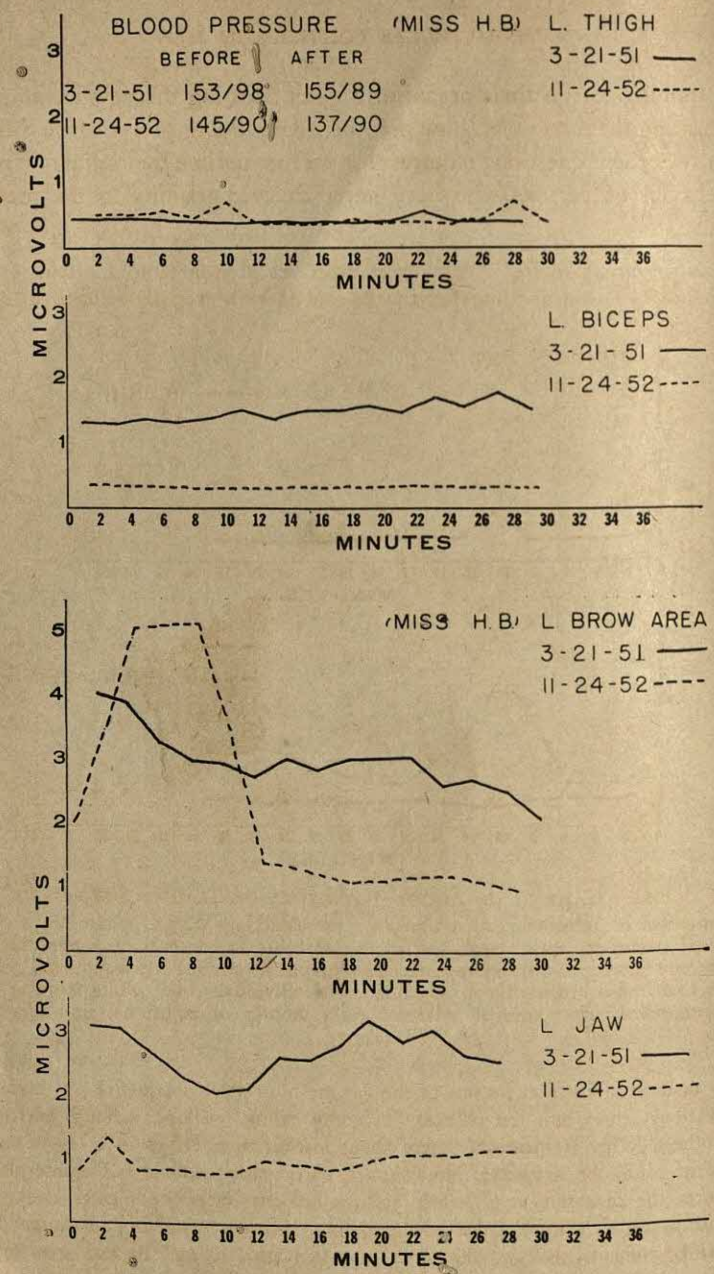


FIG. 3. SAMPLES OF ACTION POTENTIALS

Action-potentials plotted against time from four muscular regions recorded simultaneously at two periods 20 mo. apart. During this interval training was continued in a woman, age 44, under diagnosis of arterial hypertension with sclerosis. In this measurement in the lying posture, lowered potentials are least marked in the thigh; most marked in the jaw and biceps region. Increased control is illustrated, attended by lessened effort and reduced blood pressure.

nical training to run their organisms? After learning the new techniques, and making them habitual, how may their daily living performance differ? To answer these questions requires that we first outline the ordinary modes of behavior of persons who have never received training in the present sense.

Usual modes of control. That almost every act of man is to a certain extent determined by instinctive impulses can be said to be universally recognized. Thus

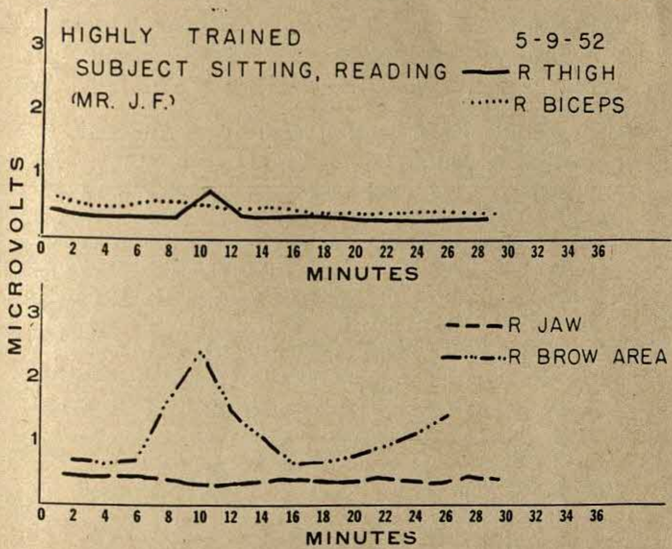


FIG. 4. SAMPLES OF ACTION POTENTIALS OF TRAINED SUBJECT
Action-potentials plotted against time in a mechanic, age 42 yr., reading in a chair. Original symptoms: vasospasm, fearfulness, inability to concentrate. Symptoms subsided only after prolonged training. The graph illustrates action-potential values lower than those ordinarily observed in most individuals, even while lying at rest. This improved effort control followed eight months of additional training while reading.

many or most of the occupations of the day are built about instinctive acts involving the various senses and the reflexes, including eating, walking, talking, exercising, and others. What is mentioned here about instinctive activities in a few sentences could of course be extended into volumes, particularly relating to the interplay of instincts, the development of habits, and the influence of environmental states.

Goal controls. With no idea of completing this sketch, we turn to another basic source of human behavior which can be called 'goal living.' By this term I refer to the ability of man to set inumerable short and long term goals for himself so that life, as he recognizes it, is largely the achieving of goals. For example, the modern worker in western cultures has goals which include bathing, dressing, and taking breakfast in the morning, later to be followed by goals variable with each

individual as the day progresses. He has variable goals for the hour, the day, the week, the month, the year, and for life. In brief, his manner of living, including his state of health at any time, is commonly determined not only by instincts and habits (and these in turn by body chemistry, the results of interplay of hereditary factors with environmental influences) but also by his innumerable goals.

There is nothing new in what has been said above about goal living, excepting the phraseology here employed. The point to be made is that goal living—the common form of everyday life—has a mechanical or physiological side which has never been realized, much less investigated.[7] If we continue our comparison of learning to run the human organism with learning to run an automobile, we can represent goal driving as something which nature has achieved in man but which automobile manufacturers have not yet attempted in cars. Perhaps at some very distant future time it may become possible for us to set a dial or other device on the dashboard stating a destination and thereupon sit back on the cushion while the car gets there. Precisely this sort of thing is what nature has accomplished in man, who proceeds to visualize or otherwise represent a goal and then to carry it out, scarcely knowing by what mechanism or at what costs of energy.

Because of ignorance concerning his own mechanisms, man, as I believe, often displays a lack of (mechanical) efficiency in various of his daily pursuits; and not knowing how to count the costs of achieving his goals (in terms of muscular sensation) he often runs his organism into hypertensive states. Thus learning to drive the organism includes the ability to count the costs (by observing states of muscular tension) and the ability to distinguish between goals and tensions of the moment. It does not, however, exclude other modes of living, whether instinctive or acquired, but modifies and supplements these in a new integration. Thus we find that persons trained according to the new principles behave as naturally as do their untrained associates, but only up to the point where the cost in their energy reserves seems to them excessive. In this respect they are practiced in making their reactions fit the situation.

There are many and various types of governing devices or controls within the living organism. Included among these varieties are chemical and hormonal substances in the tissues or in the circulation which stimulate or limit specific operations within the organism. The brain and nervous system are organized largely as an extremely complicated network of such devices, operating by the passage of electrochemical waves at speeds on the order of 5 to 100 m. per sec. These devices of control are, as Hughlings Jackson pointed out, so arranged in various levels that now we may appropriately speak of certain 'controls of controls.' Included among the latter types of controls are those which recently have been entitled 'cybernetics.' The incomprehensibly vast array of profitable researches by investigators in every branch of physiology have disclosed how many and various are the types of internal controls which engage at every instant of the activity

---

[7] Psychologists have recognized goals clearly, beginning with the epoch-making studies of Watt, Ach, Messer, and others of the Würzberg School. Their terminology has differed; namely, they used the expression *Aufgabe* where I employ the word *goal*. Furthermore, they did not conceive the organism as being driven in terms of *Aufgabe* but spoke only of determining tendencies exciting subsequent mental activities.

of the human and other animal organism. Their findings furnish a firm foundation for the position herein set forth and lend it meaning.

Learning goal controls. What we here discuss, however, are not cybernetics nor the innumerable nervous, hormonal, chemical, and other devices of control of specific activities within the organism, but rather how individuals can learn to operate their muscular mechanisms to better advantage. This is a frequent question of patients and normal individuals.

Teaching methods. The teaching methods are a further development of those employed in the cultivation of progressive relaxation, previously described. It is doubtful that anyone can fully appreciate them from a written description, since direct experience as in any other laboratory procedure is essential. For initial instruction, the subject, who has been lying on his back with eyes closed, bends his left hand back at the wrist, carefully noting the sensory experience in the wrist and volar surface of the forearm, known as 'strain sensation.' When he successfully distinguishes between these strains and the sensations in the active extensor muscles, it is tentatively agreed to call the latter experience the 'control' sensation. He can confirm the propriety of this nomenclature, in his own experience at least, by noting what he does if he exerts increasing power, as against the retarding hand of the instructor, restricting his arm movement at the wrist. As he performs these acts, the action-potential patterns may be set before him on an oscilloscope screen. If electrodes have been inserted into the extensor muscles involved, he can observe the evidences of his muscular contraction or effort and can note also to what extent he succeeds in relaxing the extensor muscles.[6] Requested to double the power which he is exerting, he can note that he uses his controls somewhat in the same manner as he directs movements of a motor car by the operation of controls on the dashboard. The analogy, to be sure, is far from complete, but it can prove a useful pedagogic comparison. The pedagogy must be in stages. So the practices included in early instruction can be only preliminary to the controlled action which is the ultimate goal. Limitations of space make it impossible to do more than illustrate some of the practices deriving from principles of control of the organism as an entirety. As indicated above, the powers of the organism can be 'on' or 'off' by direct control, a function distinguishable from so-called 'association,' well known to psychologists, and from 'conditioned and unconditioned reflexes,' well known also to physiologists.

After beginning as indicated above, the individual can be instructed similarly in the control of the 'on-' and 'off'-action of each of the principle muscle groups of his organism. Such instruction may be briefly given in a course of 20 1-hr. periods or more fully in 5 to 10 times this number, as occasion requires. In addition, as in learning any other skill, the individual needs to practice by himself. It is requisite that he be taught to distinguish the 'control sensation' from other types of sensation; that he learn to detect this 'control' as an element present in the subjective experience of mental activities, including imagination, recollection and emotion; and that he acquire the habit of using this control in his daily living. To this end he needs to be trained to observe his own sensational processes when he is facing daily problems and making adjustments to environment. It is imperative that he become accustomed to distinguish the controls (which he is learning to observe as sensations but also to

use practically) from the problem which he tries to solve. The instruction may be worded: "Distinguish between issue and tension-attitude!" Thus he may progress to the point where he recognizes that even fears and anxiety states include a measure of his own efforts to meet the environment and accordingly can be rendered more subject to daily control.

As indicated above the direct control of neuromuscular power can be rendered graphic by means of an appropriate oscilloscope. These changes can be followed objectively by the subject and the instructor at the same time. Neuromuscular television requires stability of the circuit employed, but if the action-potentials occur with grossly visible movements, they are readily recorded and no extraordinarily sensitive apparatus is necessary. If the subject, however, is to observe and the operator is to determine the onset and the decline of the control of imagination, of recall, and of other types of mental activities, the voltage sensitivity needs to be on the order of several centimeters for a single microvolt applied across the input of the amplifier. Prerequisite are various control tests which have been described previously.

Traditional psychology deals with the association of ideas or of experiences, while the Pavlovian physiology approaches the learning processes more objectively, that is, without recourse to subjective reports from the human subject. In the instance of animal subjects, such reports of course are not obtainable. Much has been learned by thousands of investigators in these scientific branches. It is necessary to emphasize, however, that when an 'association' has been formed or when a conditioned reflex has been established, the result is a more or less *automatic* determination of certain neuromuscular activity upon the occurrence of appropriate stimulation. This result derives from corresponding modifications (presumably to some extent structural) in connections within the nervous system. The ability of the organism to direct its own activities is apparently innate, at least in pristine form. This ability, I have shown, can be further cultivated to an extraordinary extent. It would be stretching the meaning of the word to call such cultivation merely a special 'conditioning.'

Cultivation of such ability depends upon (a) sufficient repetition of observation of the control process by the subject and (b) sufficient repetition of the act of control in a positive or negative direction in each principle neuromuscular locale. Such repetition chiefly in the negative direction has been previously described and analyzed from this Laboratory under the heading, *progressive relaxation.* The physiology involved concerns relationships between functions of the cerebrospinal and the vegetative nervous systems.

Results of training. In the trained subject, a distinctly different type of reaction to environment becomes possible. His responses become less determined by fixed associations or habitual conditioning. His emotions become subject to a certain measure of control. He can work with less neuromuscular tension. He can be taught to distinguish between issue and attitude—between the problems which he faces and his own neuromuscular actions. The latter constitute life while the former come to be regarded as the abstraction which they really are.

Thus efforts are freed to operate in the real interest of the organism and there can be greater efficiency, which implies less pathological wear and tear. This is because effort on the part of the organism, *i.e.,* neuromuscular activity, evidently corresponds with usage of any instrument. Notwithstanding the amazing measures of tissue restoration and repair characteristic in the organism, excessive effort tends in the long run toward organic disruption or senile change.

The life of man consists largely of efforts and sub-efforts. We do not yet know how the effects of daily efforts are related to the phenomena currently known as 'stress.' Much valuable information has been gained about glandular and steroid chemical responses upon certain types of change in the environment. It should be noted, however, that the 'alarm reaction' to stress is an unusual rather than an everyday occurrence in man; while the exertion of effort occurs every moment and constitutes much of his life.

The vast expansion of the medical sciences together with intensive specialization have resulted in neglect of the physiology of effortful adaptation. Yet we know that every muscular contraction requires oxygen and glycogen for the combustive processes and requires equally the removal of waste products, including carbon dioxide, lactic acid, pyruvic acid, and mineral ash. These requirements can be satisfied only through appropriately increased operations of the cardiovascular apparatus. Thus a life of excessive effort and strain can be expected to show overactivity of the cardiovascular system in one form or another, with possible ultimate pathology.

In our approaches to hypertensive disorders, therefore, it does not appear thorough to rely alone upon barbiturates and other measures of sedation, including psychotherapy, or upon other medications such as cyanates, sympatholytic, and ganglionic blocking drugs, or even alone upon palliative measures of salt restriction or of surgery. From what we are beginning to learn about the efforts of man in his daily adjustments to his environment and their long-time effects when in excess, it would appear that he is subject to a variety of ills which can be termed 'diseases of self-activation.' Medication, surgery, psychotherapy may in different ways alleviate the symptomatic manifestations of such diseases but without penetrating to the source of the mischief, if this really lies in waste of human efforts. Suggestions that the patient engage in hobbies, or move to another environment, or change his philosophy notoriously have failed to remove the basic cause. Whether pathological alterations have become irreversible or not, technical advances now provide measures of teaching the patient to avoid waste of effort while meeting his daily obligations.

Man has lived largely in terms of his instincts and emotions, but, less often, his reason. Hitherto, the 'life of reason' has been universally regarded as the most effective mode. The measures herein outlined differ from previous views not in being less rational but in being practical and clinical, rather than merely philosophical or popular. Man, we suggest, has failed to distinguish the situations and the issues which he faces from his own activational patterns. Only when he becomes able to recognize his own psychophysiological patterns and to distinguish them from environmental problems, does the door open for a change in his living. It it is valid and useful to regard man as thus potentially a self-driving unit provided with

controls as well as effectors in the neuromuscular system, which he can learn to operate for his own individual purposes in adjusting to environment, why has this fact been overlooked hitherto? Among assignable reasons, one stands out clearly as follows: Man is accustomed to running the controls of automobiles, planes and other constructed devices which lack a feature that distinguishes his own nervous system; namely, circuits such that when a particular control is repeatedly applied under repeatedly similar conditions, it tends to become semi-automatic in operation. To visualize what such a device might be, we can think of an automobile in the late thirties provided with a hand-operated choke to control the richness of the gasoline mixture on cold days; if repeated use of this device by hand tended to make it become operative semi-automatically, we should witness the working of mechanisms in the automobile parallel with those which underlie habit formation and conditioned reflexes in man. However, no such device has ever been developed by man and employed in his mechanical contrivances. It would be a totally new experience if after operating a control by hand repeatedly, the control began to work by itself. Because of this significant difference between the biological and the man-made controls and their operation, we assume, the possibility of man learning to drive his own organism via technical instructions has been disregarded in medicine and in the humanities.

Effective effort-control means better adaptation to environment; it can tend toward greater efficiency in every occupation of man. Activity, including reflection, can make less demands upon the cardiovascular apparatus. The blood pressure commonly tends to be lowered, not only in normal subjects but also in many hypertensives. Theoretically the incidence of coronary heart disease should be significantly diminished in groups of persons properly trained in control methods.[8] Likewise, as previously set forth, the incidence of spastic states of the alimentary tract, including those which apparently lead in the direction of peptic ulcer, should be diminished. Among other objectives, according to the principles herein outlined, it seems possible that in an adequately trained populace, the present high death rate from cardiovascular disease could be more effectively controlled. As in other divisions, preventive medicine can prove more important for social welfare than even the effective treatment of diseases once incurred. In preventive medicine, autokinetic methods should promise no 'cure' to the individual who seeks specialized education, any more than the student who registers in school is promised success in his education by the instructors. If, when, and as methods of effort education can be made available to the general populace, prevention and to some extent cure of the diseases of self-activation may become possible; but only provided that the responsibility for carrying out instructions be placed where it belongs; namely, squarely on the individual himself.

[8] Jacobson, Principles underlying coronary heart disease, *Cardiologia,* 26, 1955, 83-102.

SUMMARY

Man has autogenous controls which to some extent operate unawaredly and instinctively but which he can learn to recognize. He can be taught definitely to operate his organism as he is taught to operate man-made instruments and semi-automatic vehicles. Physiological controls are available which in many respects correspond, for example, with dashboard controls in the automobile, while muscle fiber contraction evidently corresponds with wheel revolution (Fig. 1). Important differences derive from the fact that in man the control mechanisms are self-operated, whereas in the automobile and most other man-made instruments, the driver is external. Evidence suggests that cultivated habitual use of controls can improve adaptation to modern environment, including personal efficiency and relations between men. Under modern pressures unawareness and consequent failure to employ these controls evidently leads to neuromusclar hyperkinesis with energy waste and apparently has been partly responsible for the most common of diseases; namely, the diseases of self-activation, including many fatigue states, neurotic conditions, peptic ulcer, irritable colon, and often hypertensive disorders including coronary heart disease. Physicians differ endlessly in their many and various forms of treatment administered in these conditions, including medicines, diet, surgery, removal from irritating environment and conventional psychotherapy, but notoriously the treatment generally proves only palliative at best and new symptoms soon set in. Evidently because these divers varieties of treatment all fail to penetrate to the basic cause, the diseases of self-activation continue in millions of instances, constituting a sort of modern plague. The evidence suggests that however large the undertaking, a program of effective preventive medicine will require technical education for the populace in physiological self-control.

VISUAL AND TACTILE DISCRIMINATIVE LEARNING IN PATIENTS WITH CEREBRAL TUMORS

By W. S. Battersby, H. P. Krieger, and M. B. Bender
Mount Sinai Hospital, New York City

The present study was concerned with three problems: (a) Do unilateral space-occupying lesions of the temporo-occipital areas, in man, produce maximal impairments on visual tasks of the type used in animal studies? (b) If impairments are found, are they maximal for visual tasks? (c) What relation exists between performance on visual discriminative learning tasks and the presence of primary sensory defects, or signs of 'general' mental impairment?

Gross defects in object recognition (agnosia) have frequently been described in patients with cerebral damage.[1] According to some observors, these agnosic symptoms are specific to a modality, reflecting the presence of a lesion in the corresponding 'association' area of the cortex.[2] Lesions of the parastriate area, for example, have been claimed to produce visual agnosia,[3] while lesions in the vicinity of the supramarginal gyrus are alleged to result in tactile agnosia.[4] Anatomical confirmation of cerebral damage in so-called agnosic patients has been rare, however, and where obtained has usually revealed massive or multiple lesions.[5] It has been pointed out, moreover, that the patients reported to have agnosia also had primary sensory deficits, and, in most instances, symptoms of mental deterioration in addition.[6] According to these observations, it appears impossible to differentiate 'associative' defects from either simple sensory loss or general mental deterioration. In short, as far as man is

* Accepted for publication November 18, 1954. This investigation was carried out under grant B-174S.D. of the U. S. Public Health Service, and was aided by the neurological research funds of the Mount Sinai Hospital of New York. During this study H. P. Krieger was a Fellow of the National Foundation for Infantile Paralysis.

[1] W. S. Duke-Elder, *Textbook of Ophthalmology, Vol. IV*, 1950, 3654-3665; F. B. Walsh, *Clinical Neuro-Ophthalmology*, 1947, 76-86.

[2] J. F. Fulton, *Howell's Textbook of Physiology*, 542-546 (15th Ed.). 1948; J. M. Nielsen, *Agnosia, Apraxia, and Aphasia: Their Value in Cerebral Localization*, 1936, 1-210.

[3] Nielsen, Unilateral cerebral dominance as related to mind blindness, *Arch. Neurol. & Psychiat.*, 38, 1937, 108-135.

[4] J. P. Evans, A study of the sensory defects resulting from excision of cerebral substance in humans, *Proc. Assoc. Res. Nerv. Ment. Dis.*, 15, 1935, 331-370; M. Critchley, *The Parietal Lobes*, 1953, 109-131. Following Critchey, the term tactile agnosia should be used to describe a *bilateral* deficit in object recognition, without any impairment in primary somato-sensory function. Strictly unilateral deficits, such as astereognosis, are usually (if not always) accompanied by defects in position sense.

[5] Henry Head, *Aphasia and Kindred Disorders of Speech*, 1926, 106-111.

[6] E. Bay, Disturbances of visual perception and their examination, *Brain*, 76, 1953, 515-550.

concerned, the location and extent of the lesions which cause agnosia, as well as the importance of other associated deficits remains controversial.

Early animal experiments showed that extensive, bilateral lesions of the temporal lobes resulted in gross deficits in the visual response to familiar objects.[7] More recent studies on the monkey have sought to define further the nature of these deficiencies, and the location and size of the necessary lesion. To achieve this goal the classical two-choice discrimination procedure has been employed. With this method, defects in the retention and relearning of visual discriminations have been demonstrated after bilateral aspiration of either the lateral[8] or ventromedial[9] surfaces of the temporal lobes, but not after extirpations in the vicinity of the hippocampus.[10] Lesions limited to the parastriate cortex, by contrast, have been without effect on visual tasks,[11] unless the lesions also included the lateral temporal lobe surface.[12] With reference to tactile discrimination, slight defects have been found in the monkey after posterior parietal lobe lesions,[13] but marked deficits occur only when the posterior temporal areas are also involved.[14] In general, bilateral temporal lesions in the monkey definitely disrupt performance on visual discrimination; tactile performance may also be involved, but presumably to a lesser extent.[15]

While the anatomical results of the animal studies are clear, the behavioral data are perplexing. Many factors, other than locus of lesion, seem to influence the behavior of the monkey with bilateral temporal lesions. Amount of post-operative training,[16] relative difficulty of task,[17] as well as the ability to form 'learning sets',[18]

---

[7] S. Brown and E. A. Schäfer, An investigation into the functions of the occipital and temporal lobes of the monkey's brain, *Phil. Trans. Roy. Soc., London,* 179B, 1888, 303-327; Heinrich Klüver and P. C. Bucy, An analysis of certain effects of bilateral temporal lobectomy in the rhesus monkey, with special reference to 'psychic blindness,' *J. Psychol.,* 5, 1938, 33-54; Preliminary analysis of functions of the temporal lobes in monkeys, *Arch. Neurol. & Psychiat.,* 42, 1939, 979-1000.

[8] K. L. Chow, Effects of partial extirpations of the posterior association cortex on visually mediated behavior in monkeys, *Comp. Psychol. Monog.,* 20, 1951, 187-217.

[9] M. Mishkin and K. H. Pribram, Visual discriminative performance following partial ablations of the temporal lobe: I. Ventral vs. lateral, *J. Comp. & Physiol. Psychol.,* 47, 1954, 14-20.

[10] Mishkin, Visual discriminative performance following partial ablations of the temporal lobe: II. Ventral surface vs. hippocampus. *J. Comp. & Physiol. Psychol.,* 47, 1954, 187-193.

[11] K. S. Lashley, The mechanism of vision: XVIII. Effects of destroying the visual "associative areas" of the monkey, *Genet. Psychol. Monog.,* 37, 1948, 107-166.

[12] Chow, Further studies on selective ablation of association cortex in relation to visually mediated behavior, *J. Comp. & Physiol. Psychol.,* 45, 1952, 109-118.

[13] T. C. Ruch, J. F. Fulton, J. W. German, Sensory discrimination in monkey, chimpanzee, and man after lesions of the parietal lobe, *Arch. Neurol. & Psychiat.* 39, 1938, 919-937.

[14] Josephine Blum, Cortical organization in somesthesis: Effects of lesions in posterior association cortex on somotosensory function in *Macaca mulatta, Comp. Psychol. Monog.* 20, 1951, 219-249; Blum, Chow, and Pribram, A behavioral analysis of the organization of the parieto-temporo-preoccipital cortex, *J. Comp. Neurol.,* 93, 1950, 53-100.

[15] Pribram, Personal communication.

[16] Chow, Conditions influencing the recovery of visual discrimination habits in monkeys following temporal neocortical ablation, *J. Comp. & Physiol. Psychol.,* 45, 1952, 430-437.

[17] Mishkin and Pribram, 1954, *op. cit.,* 14-20; Mishkin and M. Hall, Discrimina-

have all been shown to alter learning scores on visual discriminative tasks. Whether these same factors also influence performance on tactile tasks is unknown because little work has been devoted to this modality. It is significant that the discrimination-testing procedure has been necessary to bring out deficits in these animal studies, since gross behavioral changes in response to familiar objects have only transiently been noted in the more recent experiments. Taken together, the diverse nature of the factors influencing postoperative performance seems to indicate that no simple unitary behavioral mechanism has been destroyed by surgery.[19]

In a recent investigation on man, visual tasks were reported to be maximally impaired in cases of unilateral temporal lobe lesion.[20] This finding was interpreted as being analogous to the results found in animals.[21] The testing procedures utilized in this investigation, however, were not similar to those of the animal studies. Where similar procedures have been employed in man, location of lesion has not been found to be a major determinant of defect. Performance on complex visual[22] and tactile[23] tasks, for example, has been shown to be impaired by any unilateral cerebral lesion, irrespective of location. These latter studies, however, were limited to World War II veterans with penetrating gunshot injuries of the cerebrum. Location of lesions in this population is never precise, nor are agnosic symptoms likely to occur in such patients, since they have recovered from a cerebral injury sustained many years prior to testing. Patients with cerebral tumors, on the other hand, are more likely to have symptoms resembling agnosia, since progressive cerebral damage and concomitant behavioral defects are produced by such lesions.[24] In tumor cases, moreover, the location of the lesion can more readily be defined in terms of standard radiographic techniques, since it is space-occupying in nature.

METHOD

Subjects. On the basis of radiographic evidence (angiography, pneumoencephalography, or ventriculography) and surgeon's notes upon craniotomy, 69 randomly chosen patients with space-occupying lesions of the cerebral hemispheres were grouped as follows: Group *F*, 22 patients with lesions involving the frontal and

---

tion along a size continuum following ablation of the inferior temporal convexity in monkeys, *J. Comp. & Physiol. Psychol.* (in press); Pribram and Mishkin, Simultaneous and successive visual discrimination by monkeys with inferotemporal lesions, *J. Comp. Physiol. Psychol.* (in press).

[18] Chow, Effects of temporal neocortical ablation on visual discrimination learning sets in monkeys, *J. Comp. & Physiol. Psychol.*, 47, 1954, 194-198; A. J. Riopelle, R. G. Alper, P. N. Strong, and H. W. Ades, Multiple discrimination and patterned string performance of normal and temporal-lobectomized monkeys, *J. Comp. & Physiol. Psychol.*, 46, 1953, 145-149.

[19] K. L. Chow and P. J. Hutt, The 'association cortex' of *Macaca mulatta:* A review of recent contributions to its anatomy and functions, *Brain*, 76, 1953, 625-677.

[20] Brenda Milner, Intellectual effects of temporal lobe damage in man, Unpublished Ph.D. thesis, McGill Univ., 1952.

[21] Milner, Intellectual function of the temporal lobes, *Psychol. Bull.*, 51, 1954, 42-62.

[22] H. L. Teuber, W. S. Battersby, M. B. Bender, Performance of complex visual tasks after cerebral lesions, *J. Nerv. & Ment. Dis.*, 114, 1951, 413-429.

[23] Josephine Semmes, S. Weinstein and H. L. Teuber, Performance on complex tactual tasks after brain injury in man: Analyses by locus of lesion, this JOURNAL, 67, 1954, 220-240.

[24] M. B. Bender, *Disorders in Perception*, 1952, 16-28.

fronto-parietal areas; Group *TO*, 18 patients with masses in the temporal and temporo-occipital areas; and Group *PO*, 21 patients with lesions of the parietal and parieto-occipital areas. Eight patients with large lesions involving the parieto-temporal, and in part frontal areas, were added to this group. Post-mortem examination confirmed the extent of brain damage in 8 of the experimental cases.[25]

Two different categories of control *Ss* were studied: Group *SC*, 12 patients with lesions limited to the spinal cord; and Group *PF*, 17 patients with lesions of the posterior fossa, or with generalized increased intracranial pressure due to other unknown causes.

Many patients in each group, including controls, were acutely ill and could be examined on only some of the procedures. In many instances the same patients could not be tested both pre- and post-operatively.[26]

Procedures: (a) Primary sensory abilities. All *Ss* received tests of color vision (isochromatic plates), extent of visual field (confrontation), and where indicated, perimetric fields for 1° white, red, and blue targets. Clinical examinations for anomalous eye movements, diplopia, and metamorphopsia were also performed. In somesthesis, the perception of touch and pin prick, two point thresholds, and adaptation-time were tested. In addition, position-sense, stereognosis for common objects held in hand, and perception of letters and numbers traced on the skin (graphesthesia) were evaluated.

(b) Mental status. The orientation of each *S* for time, place, and person, plus his memory for recent and distant events, was assessed by means of standard questions. His ability to perform serial subtractions, to follow simple and complex verbal commands, to name and describe the utility of common-place objects, as well as any tendencies for expressive or receptive aphasia were noted in detail. In many cases, injection of Amytal Sodium (R) was employed as an aid in determining the mental status of the patient.[27]

(c) Discriminative learning. Two parallel tests of discriminations were administered; one visual, the other tactile. Each set consisted of five series, which, in the order given, involved either form, size, intensity (roughness or brightness) or conditional discriminations. Series 1 was a triangle vs. circle (form) discrimination. Series 2 was composed of a larger vs. smaller discrimination of intermixed triangles and circles. Series 3 was a multiple-cue discrimination involving a combination of intensity and form discriminanda. For example, in the visual series, the darker of two triangles, and the lighter of two circles were positive, while on the tactual test, the smoother triangle and rougher circle were consistently rewarded. Series 4,

---

[25] With but one exception (Case No. 26, *vide infra.*), patients with metastatic disease were excluded from this study unless post-mortem examination confirmed the extent and location of cerebral damage.

[26] All operative procedures were carried out by members of the Department of Neurosurgery, Mt. Sinai Hospital, under the direction of Dr. L. M. Davidoff. We express our appreciation to this department for their active coöperation during the course of this study.

[27] We are indebted to Drs. E. A. Weinstein and R. L. Kahn for making these data available. For details on the use of Amytal Sodium in determining mental status see E. A. Weinstein, R. L. Kahn, L. A. Sugarman, and L. Linn, Diagnostic use of Amobarbital Sodium (Amytal Sodium) in organic brain disease, *Amer. J. Psychiat.*, 109, 1953, 889-894.

another multiple cue discrimination, involved both relative size and intensity, *i.e.* the larger of two dark (or smooth) stimuli, the smaller of two light (or rough), irrespective of form. Series 5 was a traditional conditional discrimination. A triangle was consistently rewarded on a horizontally striped background while a circle was positive if the stripes were vertical.

Following the procedure used in animal studies, the stimuli were presented in pairs, by the projection of slides for the visual series and by means of blocks behind an opaque curtain for the tactual series. *S* was told that only one of the two figures would be 'correct,' and, on the basis of what he saw (or felt), to keep trying to get as many 'correct' responses as possible on subsequent trials. In Series 1 only, the stimulus to be consistently rewarded was pointed out to *S* on the initial trial to make sure that the procedure was understood. On subsequent trials *S* was told only 'correct' or 'incorrect,' depending upon his choice. Within each series the left-right position of the consistently rewarded stimulus was randomized on the basis of a Gellerman order.

The results for Series 1–4 were scored in term of the number of errors made in reaching a criterion of 10 consecutive correct responses (up to a maximum of 30 consecutive trials); 12 consecutive correct responses (to a maximum of 48 consecutive trials) was the criterion used on Series 5. We were able to complete Series 1–4 on almost all cases. Series 5, however, was restricted to those patients who had solved either Series 3 or 4, or both. It is interesting that Series 5 produced such marked 'catastrophic' behavior in patients with noticeable defects that clinical management dictated that this series be restricted to patients with minimal impairments.[28]

To obtain equivalent series for post-operative testing, the brightness (or roughness) of figure and background were interchanged, and the previously positive stimulus-figure was made negative; *i.e.* never reinforced. These equivalent forms, as well as the order of the visual and tactile sets were randomized over the group of subjects tested.

RESULTS

Quantitative. For all groups the mean error scores on Series 3 and 4 were much larger than on Series 1 and 2. Differences between groups on Series 1 and 2 would have involved values of less than one—less than the accuracy of measurement. To obtain a more reliable index of individual performance, therefore, the total errors per *S* were calculated for Series 1–4. Data for Series 5 were treated separately since only selected *Ss* received this test. Table I presents the group means and standard deviations for the total error scores for Series 1–4, and Table II the same means and standard deviations for the errors made in learning Series 5. Pre- and post-operative data are given for both visual and tactile series.

The data show no marked differences among the experimental groups with the possible exception of the total error score for Group *PO* on the tactile series. The *SD*, however, show a high degree of inter-individual

[28] Kurt Goldstein, *After-Effects of Brain Injuries in War*, 1942, 1-244.

variability within each of the experimental groups. With reference to the controls, only slight differences are found between Groups *SC* and *PF*, although both of these groups appear to make less errors than the experimental *S*s. Again, however, inter-individual variability is substantial.

TABLE I

SUMMARY OF ERROR SCORES: SERIES 1–4

Group	Pre-operative						Post-operative					
	visual			tactile			visual			tactile		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
F	13	20.9	9.1	14	25.4	10.0	13	25.5	9.9	12	20.5	11.7
PO	13	22.2	7.6	8	30.0	17.1	16	20.5	9.1	11	27.3	8.8
TO	12	21.3	8.2	12	19.3	14.7	11	21.1	13.3	7	24.7	10.6
Total	38	21.4	8.3	34	23.9	14.1	40	22.3	11.0	30	23.8	11.0
SC	8	14.4	6.0	6	15.2	8.3	5	16.0	6.7	6	18.6	7.1
PF	16	18.7	8.7	14	15.9	8.5	12	13.6	10.2	9	15.8	11.2
Total	24	17.2	8.2	20	15.7	8.5	17	14.3	9.4	15	16.8	10.0

The results of the statistical analyses of differences between groups confirm the impressions gained by inspection of Tables I and II. The mean difference *t*- and *P*-values were calculated for the visual and tactile tests and for both pre- and post-operative conditions. In general, no significant differences were found between any of the experimental groups in terms of total error scores (Series 1–4). Moreover, no

TABLE II

SUMMARY OF ERROR SCORES: SERIES 5

Group	Pre-operative						Post-operative					
	visual			tactile			visual			tactile		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
F	11	19.6	9.6	8	16.8	12.5	10	16.8	11.0	9	12.3	12.5
PO	11	20.9	7.2	3	25.3	2.6	15	13.7	11.0	5	15.2	11.8
TO	10	15.1	6.4	7	15.9	13.0	8	15.6	8.8	2	15.0	15.0
Total	32	18.7	9.4	18	17.8	12.1	33	15.1	10.6	16	13.6	12.7
SC	6	14.5	10.6	4	6.5	9.6	5	18.4	0.8	4	18.8	10.3
PF	16	15.4	9.5	13	14.0	13.3	12	7.3	8.0	9	8.1	10.4
Total	22	15.1	9.8	17	12.2	12.9	17	10.6	8.4	13	11.4	11.8

significant differences exist between the two classes of controls, Groups *PF* and *SC*. By contrast, significant differences between combined control and combined experimental groups can be demonstrated on visual, and on tactile total error scores for both pre- and post-operative conditions. These differences are shown in Table III. On Series 5, no significant differences were demonstrable between any of the groups, experimental or control, including combined experimental vs. combined control. The totals are shown in Table III. Since the *S*s receiving this series were

selected on the basis of their performance on Series 3 and 4, however, the lack of significant differences on this particular series is not surprising.

Additional *t*-tests were performed within each group between the total error scores on the visual as compared to the tactile series. None of these comparisons achieved statistical significance as can be seen from Table IV. These negative findings imply that both the visual and tactile series were sampling some common

TABLE III

INTERGROUP COMPARISONS OF ERROR SCORES: TOTALS OF EXPERIMENTAL AND CONTROL GROUPS. (Mean difference is experimental minus control)

Series	Pre-operative visual $M_{Diff.}$	t	P	Pre-operative tactile $M_{Diff.}$	t	P	Post-operative visual $M_{Diff.}$	t	P	Post-operative tactile $M_{Diff.}$	t	P
1–4	4.2	1.91	>.05	8.2	2.53	<.02	8.0	2.55	>.01	7.0	1.96	<.05
5	3.6	1.33	.20	5.6	1.30	.20	4.5	1.50	.20	2.2	0.47	.70

ability; an ability not specific for the modality through which the tests were administered. To evaluate this possibility, the Pearsonian correlations between total errors scores on visual and tactile series were computed separately for the combined experimental and combined control groups. The pre-operative correlations for these experimental and control groups were +0.45 (significant at the 1-% level of confidence), and +0.70 (significant at better than 1-% level of confidence, respectively.

TABLE IV

INTERMODALITY COMPARISONS OF ERROR SCORES: SERIES 1–4 (Visual Minus Tactile Mean Differences per Group)

Group	Pre-operative $M_{Diff.}$	t	P	Post-operative $M_{Diff.}$	t	P
F	−4.5	1.15	.30	5.0	1.11	.30
PO	−7.8	1.26	.30	−6.8	1.78	.10
TO	2.0	0.39	.70	−3.6	0.54	.60
SC	−0.8	0.19	.90	−2.6	0.54	.70
PF	2.8	0.36	.80	−2.2	0.45	.70
Total Experimental	−2.5	0.90	.40	−1.5	0.54	.60
Total Control	1.5	0.59	.60	−2.5	0.69	.50

Post-operatively, the correlations were +0.55 for the experimental group (significant at better than the 1-% level), and +0.65 for the controls (2-% level).

In summary, the group data indicate that unilateral temporo-occipital lesions in man are no more disruptive of visual discriminative learning than lesions elsewhere. Moreover, defects in discriminative learning, when found, do not appear to be restricted to one modality. Finally, differences between brain injured and control *S*s can be demonstrated only on a statistical basis, since the groups overlap greatly in their performance. This

can be seen from Figs. 1 and 2 which present the cumulative relative frequencies of the visual and tactile error-scores for pre- and post-operative conditions respectively.

Qualitative Observations. Results so far suggest that performance on the discrimination tasks used in this study was dependent upon some sort of general ability, rather than upon factors specific for each modality. Table V presents for the combined experimental group the relative incidence of visual, somato-sensory, phasic, and mental deficits, as a function

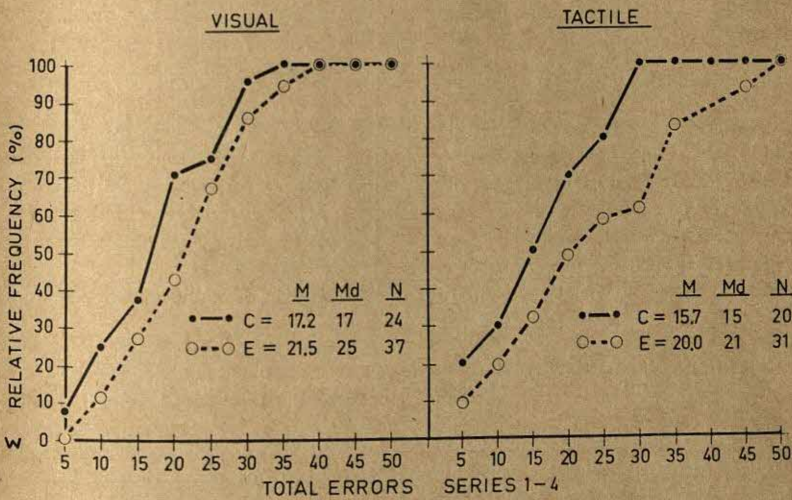


FIG. 1. CUMULATIVE DISTRIBUTION OF THE SCORES MADE PRE-OPERATIVELY EXPERIMENTAL GROUPS, BROKEN LINE; CONTROL GROUPS, SOLID LINE.

of the total error scores on the visual learning series administered pre-operatively. It is readily seen that as the total error-score increases, the relative incidence of mental deficits does the same. The relative incidence of somato-sensory, phasic, or visual defects, on the other hand, seems to bear little relation to the total error-score on the learning series. Numerous observations made in the course of this study support the hypothesis that the general mental status of the patient was the most important determinant of performance on the discriminative learning tasks.

Case No. 22, a factory foreman, 53 yr. old, underwent craniotomy for partial removal of a left temporo-occipital glioblastoma. Both pre- and post-operatively this man had an incomplete right homonymous hemianopia, questionable right somato-sensory defects, a marked disorientation for time, place, and person, and severe dyscalculia with difficulties in identifying numbers. His pre-operative total error score (40 errors on visual, 32 on tactile series), as well as his post-operative per-

formance (40 errors on visual, 31 on tactile tests) was noticeably inferior to that of the average brain-damaged patient (as can be seen from Figs. 1 and 2). By contrast, Case No. 4, a 24 yr.-old college student, was operated upon for partial resection of a right temporo-occipital sarcoma. Both pre- and post-operatively he did

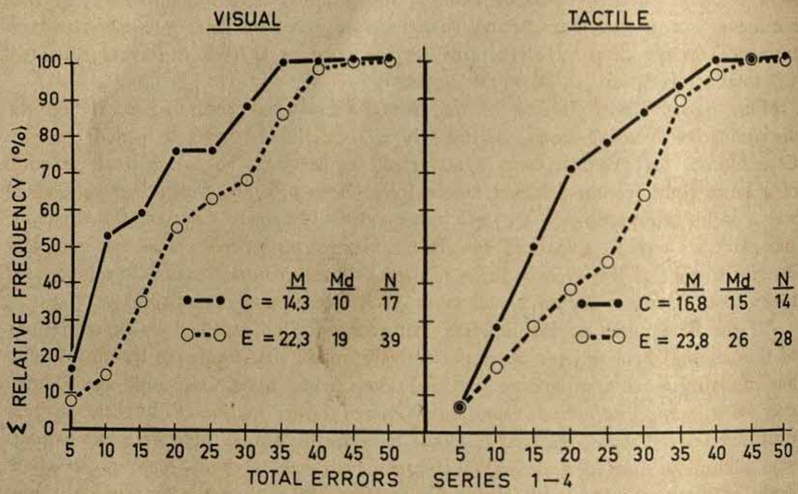


FIG. 2. CUMULATIVE DISTRIBUTION OF THE SCORES MADE POST-OPERATIVELY EXPERIMENTAL GROUPS, BROKEN LINE; CONTROL GROUPS, SOLID LINE.

much better than the average control (pre-operatively, 17 total errors on visual, 5 on tactile test; post-operatively 14 and 16 total errors, respectively). This patient had a left homonymous hemianopia, and a syndrome of left hemisensory deficit, including defects in stereognosis and position sense. He was completely

TABLE V

RELATIVE INCIDENCE OF NEUROLOGICAL DEFICITS BY TOTAL ERROR SCORE ON VISUAL DISCRIMINATIONS (PRE-OPERATIVE)

Total errors	Frequency of Brain Tumor Cases	Relative incidence of deficits (%)			
		visual	somato-sensori	phasic	mental
0–9	4	25	25	0	25
10–19	13	38	54	15	15
20–29	15	40	60	40	60
30 or more	5	41	20	41	100
Total	37	38	49	27	46

oriented during the period of testing and never showed any signs of disorientation, dyscalculia or memory loss. The location of the lesion was confirmed on post-mortem examination.

Although the lesion in the two foregoing cases were similar, their age and socio-economic background differed widely. In other cases, however, these factors

by themselves have seemed of little importance. This is illustrated by Case No. 16, a 46 yr.-old immigrant housewife, who underwent a partial removal of a metastatic carcinoma from the left temporal lobe. She died soon after operation and the location and extent of lesion was verified at autopsy. Clinically, this patient had showed a right homonymous hemianopia and minimal right somato-sensory defects, but never manifested any mental changes during testing. Her pre-operative performance on the discriminative tasks was as good as that of the average control (22 total errors on visual, 20 on tactile tests).

Many patients with lesions of the frontal lobes also performed poorly on the discriminative learning tasks, particularly if they showed marked mental changes. Case No. 5, for example, was a 60 yr.-old housewife who had a total resection of a large right frontal calcified meningioma. Both pre- and post-operatively, there was a mild hemiparesis of the left upper extremity, severe disorientation for time and place, as well as a marked dyscalculia. Her learning performance was markedly inferior to that of the average brain injured *S* (total error scores of 32 on the visual series, pre-operative; 35 on visual and 31 on tactile, post-operatively).

All the data given so far indicate that anatomically confirmed unilateral lesions of the cerebral hemispheres do not selectively impair discriminative learning within one modality when comparisons are made on a group basis. One could argue, however, that some *individual* cases might have shown marked differences in performance on visual as compared to tactile tests. In this study, there were two patients who definitely showed a defect much greater in one modality than in the other. Both of these patients made only chance error scores on even the simplest form discrimination (Series 1) of the tactile series. Both of them, however, also had marked mental changes (disorientation, dyscalculia, and memory loss), and in addition, did worse than the majority of brain damaged *Ss* in learning the visual series (total error scores of 34 and 35). Even more confusing is the fact that the pathological anatomy in these two cases was somewhat different. One patient had a left occipital pole resection for removal of a meningioma; the other had a large glioblastoma of the right parietal area, confirmed upon autopsy. Both of these lesions were in the posterior brain substance, yet their exact locations and etiologies were quite different. Despite this fact the behavioral deficiencies of these two patients superficially appeared similar.

Discussion

Visual discriminatory learning and location of cerebral lesion. The results obtained in this investigation indicate that the location of a unilateral space-occupying lesion in man is not a major factor in producing decrements in visual discriminative learning. This negative finding cannot be attributed to diffuse pressure effects of cerebral tumors, since patients with only generalized increased cerebral pressure (group *PF*) did significantly better than those with space-occupying lesions. In general, our results confirm and extend those of earlier studies on the effects of penetrating shrapnel wounds of the cerebrum in man. It has been shown, for example, that gunshot injuries of the frontal, parietal, temporal, or occipital areas

interfere equally with performance on visual discriminatory tasks.[29] Present results are also compatible with those of our earlier study, which demonstrated that tumors of these same areas were equally disruptive on tasks requiring complex visual perception.[30]

The results of the present study differ, however, from those reported for a series of patients receiving cerebral excisions for removal of epileptogenic scar tissue.[31] Milner found that unilateral temporal lobe lesions produced maximal deficits on some visual tests (McGill Picture Anomalies and Wechsler-Bellevue Picture Arrangement), but not on others (Wechsler-Bellevue Picture Completion, Benton Visual Retention, and the Pictorial Analogies from the California Tests of Mental Maturity). Differences between patients with frontal and temporal lobe lesions were small, however, and, as stated, were obtained on only some of the visual tests used. This finding indicates either that only certain visual tasks depend upon the temporal lobe, or, more parsimoniously, that the differences are attributable to some factor other than the visual nature of the tests. Patients with parietal lesions were not included in her study; thus, the specificity of the deficit in behavior with lesions in the temporal was not demonstrated. In our study there were five cases with unilateral temporal lesions, confirmed upon autopsy, who performed as well as the average control. These negative cases argue strongly against attributing visual integrative functions exclusively to the temporal lobes of man.

It may be argued that bilateral temporal lesions are necessary to produce maximal deficits in man. It is true that all the experiments to date have emphasized that bilateral lesions are essential in the monkey to produce deficits on visual discrimination. Bilateral and delimited tumors involving the lateral surfaces of the cerebral hemispheres are extremely rare in man. In our study there was only one case with evidence of bilateral temporal lesions.[32] This patient, Case No. 26, was a 42 yr.-old righthanded copywriter, who received a partial resection of the right temporal lobe for removal of a metastatic carcinoma. Three months later, she was readmitted to the hospital with neurological signs referable to the left fronto-temporo-parietal area (aphasia, paresis of the right upper extremity and somatosensori deficits in the right hand). She was completely oriented and coöperative until a few days prior to death. Confirmation of the location of the lesions was unfortunately not obtained in autopsy. This patient received one of the highest scores of all patients tested, controls included, on both visual and tactile series (5 total errors on visual, 5 on tactile pre-operatively; 6 total errors on visual, 2 on tactile post-operatively). Presumed bilateral temporal lobe damage in this case had little effect, apparently, on the ability to learn the type of discriminations used in this study.

This finding is compatible with observations recently made on the effects of bilateral resections of the temporal lobe for relief from psychoses.[33] Although formal testing

---

[29] Teuber, Battersby, and Bender, *op. cit.*, 413-429.

[30] Battersby, H. P. Krieger, M. Pollack, and Bender, Figure-ground discrimination and the "abstract attitude" in patients with cerebral neoplasms, *Arch. Neurol. & Psychiat.*, 70, 1953, 1-10.

[31] Milner, 1954, *op. cit.*, 42-62.

[32] We are grateful to Dr. Sidney W. Gross for allowing us to study this private patient in detail.

[33] S. Obrador, Temporal lobotomy, *J. Neuropathol. & Exper. Neurol.*, 6, 1947,

procedures have not been employed on these patients, no gross changes in their response to visual objects have been noted. More work on this type of material may be of value in evaluating further the behavioral effects of bilateral temporal lobe lesions in man.

Modality of deficit after temporal lesion. We have demonstrated that performance on tactile learning is usually impaired by the same lesions that interfere with visual discriminatory learning; in neither case does the location of the lesion appear to be crucial for man. A recent study on the effects of cerebral gunshot injuries has indicated that defects in tactile discriminatory learning are produced by any unilateral cerebral lesions, irrespective of location.[34] Our results confirm and extend these findings, and in addition support the contention of a general deficit, rather than one specific for any modality.

Once more, our interpretation appears to be inconsistent with the results obtained by Milner.[35] Her conclusion was based on the fact that patients with temporal lesions did worse than those with frontal lesions on visual tests, while no significant differences were demonstrated on tasks presumable more tactile in nature (the block design of the Wechsler-Bellevue, and the modified Sequin form board). One can debate, however, the advisability of defining these latter tasks as tests of tactual ability; block design and form board tasks would appear to depend, at least in part, upon some degree of spatial perception. In any event, as far as the form board is concerned, performance has already been shown to be impaired by unilateral cerebral lesions of the frontal, temporal, parietal, or occipital areas.[36]

In general, it appears hazardous to label a defect in complex performance as specific to a modality without taking into account the multiplicity of factors that can determine test-behavior. It is true that deficits much greater in one modality than in another may be seen in patients with cerebral lesions; but more significant, perhaps, is the finding that these cases also show noticeable deficits in other modalities when adequate testing methods were utilized. This is not necessarily at variance with animal studies, since the latter have not demonstrated that uniquely visual deficits are found after bilateral lesions of the temporal lobes. In fact, the many factors that influence the behavior of monkeys with bi-temporal lesions suggests that it is not a specific visual deficit that has been created by surgery.[37] Klüver has stressed the fact that his monkeys with 'psychic blindness' were not able to differentiate objects on the basis of either visual or tactile cues.[38] His monkeys discriminated

---

185-193; W. B. Scoville, R. H. Dunsmore, W. T. Liberson, C. E. Henry, and A. Pepe, Observations on medial temporal lobectomy and uncotomy in the treatment of psychotic states, *Res. Pub. Assoc. Nerv. & Ment. Dis.*, 31, 1953, 347-373.

[34] Semmes, Weinstein, and Teuber, *op. cit.*, this JOURNAL, 67, 1954, 220-240.

[35] Milner, 1954, *op. cit.*, 42-62.

[36] Teuber and Weinstein, Performance on a formboard-task after penetrating brain injury, *J. Psychol.*, 38, 1954, 177-190.

[37] Lashley, In search of the engram, in *Physiological Mechanisms in Animal Behavior*, 1950, 454-482.

[38] Klüver and Bucy, 1938, *op. cit.*, 33-54.

food from inedible objects solely on the basis of oral exploration. In summary, experiments on animals and studies of men with brain injuries seem to agree in that the nature of the observed deficits are of greater complexity than one would predict from theories of 'associative,' or simple sensory loss after cerebral damage.[39]

Summary

In this study 69 patients with space-occupying lesions in various areas of the cerebral hemispheres were tested for their ability to learn visual and tactile discriminations of the type used in recent animal experiments. The learning scores of patients with temporo-occipital, parieto-occipital, or frontal lesions did not differ significantly, although as a group they were significantly inferior to patients with infra-tentorial lesions.

No significant differences between visual and tactile learning scores could be demonstrated within any of the groups. In *Ss* with hemispheric lesions, learning performance appeared to be related to general mental status, and was relatively unaffected by the presence of primary sensory defects. These results are evaluated in light of recent theories of an integrative focus in the temporal lobes for the elaboration of visual impulses.

---

[39] Bender and Teuber, Psychopathology of vision, in *Progress in Neurology and Psychiatry*, 1949, 163-172; Semmes, Agnosia in animal and man, *Psychol. Rev.*, 60, 1953, 140-147.

PATTERNS OF EXPERIENCE AND THE CONSTANCY OF AN INDIFFERENCE POINT FOR PERCEIVED WEIGHT *

By William Bevan and Charles L. Darby, Emory University

The inadequacy of the Weber-Fechner law as a predictive tool has become increasingly clear in recent years. Should the traditional indictments be effectively answered, this evaluation would still hold, for the Fechner formulation makes no provision for assessing the contribution of context to the response.[1] Context is derived from two sorts of variable: the properties of the present stimulus-complex and past experience. Helson's demonstration that color responses vary with changes in background reflectance and Johnson's that judgments of pitch are affected by immediately previous training with frequencies differing from those of the test-range illustrate the importance of context for discrimination tasks of even a simple sort.[2]

Context effects are reflected by shifts in the indifference point. Johnson refers to this value, obtained by the method of absolute judgment, as the category limen between "highness" and "lowness" and defines it as "the geometric mean of the frequencies of the stimuli to which . . . [the subject] has been exposed":[3]

$$\log x_o = \Sigma \log X_i/N.$$

This is the same as Helson's specification of the adaptation-level for lifted-weights with the same method, except that Helson includes a correction for time-order and the size of the stimulus-interval:[4]

$$\log (AL + .75d) = \quad \log X_i/N.$$

Response magnitude is thus represented as a function of the difference between a present stimulus-process and an internal standard evolved from a combination of previously occurring relevant stimulus processes.

* Accepted for publication February 11, 1955.

1 E. G. Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 44-45.

2 Harry Helson, Fundamental problems in color vision: I. The principle governing changes in hue, saturation, and brightness of non-selective samples in chromatic illumination, *J. Exper. Psychol.*, 23, 1938, 439-476; Helson and V. B. Jeffers, Fundamental problems in color vision: II. Hue, brightness, and saturation of selective samples in chromatic illumination, *ibid.*, 26, 1940, 1-27; D. M. Johnson, Learning function for a change in the scale of judgment, *ibid.*, 39, 1949, 851-860.

3 Johnson, Generalization of a reference scale for judging pitch, *J. Exper. Psychol.*, 39, 1949, 317.

4 Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this Journal, 60, 1947, 1-30.

It now becomes necessary to ask under what conditions a stimulus is relevant. It is reasonable to expect that not all stimuli having impinged upon the organism affect the standard. In the case of lifted-weights, for example, one may conceive of weight experiences so widely in magnitude or so far removed in time those associated with the task at hand that they exert no influence at all upon the latter. This conclusion gains some support from experimental results.[5]

The criteria for distinguishing relevant from irrelevant stimuli are not provided by the Helson and Johnson formulae. This is primarily an emperical problem of establishing limits and identifying variables. The psychophysical situations utilized by Helson and Johnson are such, however, that we may promise the crucial variables to be limited in number. The formulae assume these to be two, reciprocally interrelated: The magnitude of the stimuli on the principal physical variable (*e.g.* frequency for pitch, weight for heaviness), and the frequency with which each stimulus-magnitude occurs. The indifference point (*IP*) assumes a particular position, not because *S* has been exposed to certain magnitudes or has had some specific amount of practice, but because he has been presented certain magnitudes, each a particular number of times.

The problem of relevance may be approached, therefore, by asking: For any given *IP*, is there a range of magnitudes within which this reciprocity is maintained? Or more precisely: Given a common core of experience, what stimulus-magnitudes may be added to experience, and in what fashion, without affecting the position of the *IP* on the dimension under scrutiny?

Since there is evidence to suggest that relevant past experience is confined to the test-session under consideration,[6] the above question may be specified as follows: If 'come core of experience' be defined as that provided by four weights, 220, 260, 300, and 340 gm., each presented three times, and supplemnted by 'additional experience' in the development of an *IP* at 280 gm., what combinations of magnitude and frequency of presentation may be utilized to define 'additional experience'?[7] Fig. 1 pro-

---

[5] S. Rogers, The anchoring of absolute judgments, *Arch. Psychol.*, 37, 1941, (No. 261), 1-42; Helson, Perception, in H. Helson, (ed.), *Theoretical Foundations of Psychology*, 1951, 382; Johnson, How a person establishes a scale for evaluating his performance, *J. Exper. Psychol.*, 36, 1946, 25-34; W. Bevan and P. Saugstad, Breadth of experience, ease of discrimination, and efficiency of generalization, *Brit. J. Psychol.*, in press; M. E. Tresselt, The influence of practice upon the formation of a scale of judgment, *J. Exper. Psychol.*, 37, 1947, 251-260; The effect of the experiences of contrasted groups upon the formation of a new scale of judgment, *J. Soc. Psychol.*, 27, 1948, 209-216.

[6] Helson, *ibid.*, 382; Johnson, *ibid.*, 30; W. Bevan and P. Saugstad, *loc. cit.*

[7] Note that Helson's formula does not predict an indifference point at the center of the 'core' series. The additional experience is required, therefore, to bring the indifference point back to the center of the 'core' series.

vides predictions of these values based on Helson's formula. It was derived by setting the left side of Helson's equation at 310 (280 + 0.75 times th estimulus-interval of the core), using as the denominator on the right side a series of $12 + n$ values ('core' + 'additional experience' trials) and solving for $\Sigma \log X_i$. The sum of the logs of the core stimuli was then subtracted from this solution, the antilog of the difference equaling the stimulus-magnitude complementing the n frequency used.

With the curve of equivalent 'additional experience' constructed, the experimenter's (*E*'s) first task is to select from it certain combinations of frequency and stimulus-magnitude, combine these with the core, and compare the resultant *IP*s with the pre-

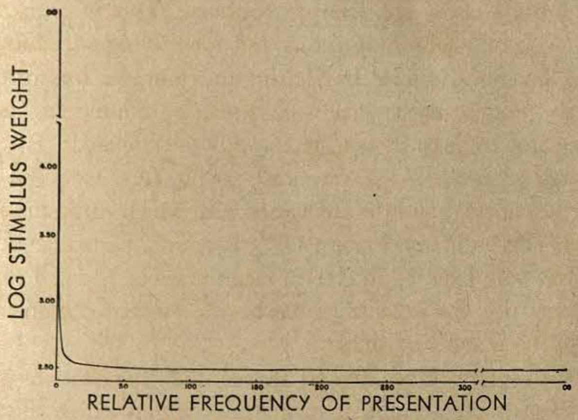


FIG. 1. PREDICTED STIMULUS AND FREQUENCY VALUES
The curve shows stimulus magnitudes that, when multiplied by frequency of presentation and combined with "core" experience, predict an indifference point of 280 gm.

dicted values. Should all points on the curve yield the same indifference point, Helson's principle would hold without exception. Should only points within some particular segment of the curve of Fig. 1 provide the predicted *IP*, limits within which the principle holds would have been established. We would also have available a model that allows, for this simple situation, designation of a set of *stimulus-equivalence relationships* without recourse to a physiological fiction or the necessity of empirically justifying each member of the set.

EXPERIMENTAL METHOD

The design has just been described. Groups of *S*s were tested with various frequency—magnitude combinations of 'additional experience' until high frequency–low magnitude and low frequency–high magnitude values were found that failed to yield the predicted *IP*. It proved necessary to introduce changes in instructions at different stages of the experiment in order to correct for the effects of instructions upon the shape of the psychophysical function.

Materials. Weights were constructed by filling plastic containers, 2 in. square and 4 in. high, with lead shot and cotton. The heaviest weight was filled with mercury. The four core weights were 220, 260, 300 and 340 gm. Six additional weights, one for each experimental group, were the stimuli for 'additional experience:' 312.1, 314.4, 1010 used in two groups, 1231 and 1950 gm.

Subjects. The *Ss* were 60 undergraduate psychology students, none of whom had previously taken part in a psychophysical experiment, assigned randomly into 6 groups of 10 *Ss* each, and tested individually in a single session of approximately 25 min. duration.

Groups. Each *S* made between 204 and 224 judgments of single stimuli with five weights randomly. The *Ss* belonged to the following groups:

Groups 1 and 3 provided the first attempt to bracket the range within which the principle of adaptation-level holds. They received rating instructions in their original form. Group 1 was given 224 trials, the core 17 times, plus 1010 gm. 20 times. Group 3 received 224 trials, the core twice plus 314.4 gm. 207 times.

Group 2 received a repetition of the stimulus-series of Group 1 but with modified instructions.

Groups 4 and 5 represent a second attempt to bracket the range. Group 4 received 221 trials, the core 17 times, plus 1231 gm. 17 times. Group 5 made 212 judgments, the core once and 312.1 gm. 200 times.

Group 6 was run to further explore the effect of using frequency-magnitude combination outside the bracketed range. It received 204 trials, th ecore 16 times plus 1950 gm. 140 times. Groups 4, 5, 6 used a second modification of the instructions.

Procedure. S and *E* sat on opposite sides of a Z-shaped screen, the former with his non-dominant hand extended through an aperture in the middle panel. Before any judgments were made, *E* described the proper technique of lifting, recited the usual precautions, and gave a demonstration of the correct manner of handling the weights.

Instructions. Groups 1 and 3 were instructed as follows in the use of Helson's nine-category rating scale:

You will be presented one weight at a time. After lifting it, you are to give a judgment by calling it very, very heavy; very heavy; heavy; medium heavy; medium, medium light; light; very light; or very, very light. If you need other categories you may use them. To help you a card with the categories typed on it is attached to the screen in front of you.[8]

Since the data of Group 1 did not yield the expected rectilinear curve, the same series was presented to Group 2 with two additional formally specified judgment categories: very, very, very, heavy; and very, very, very light.

The modified instructions produced the expected results. Since it was desirable, however, to have a set of instructions that would not require change each time the range of magnitudes in the stimulus-series was increased, further modification of the instructions was considered. Thus far, they tended to stress the finite number of units of which the judgmental scale was to consist. Hence they may have conveyed the impression that responses should be more or less confined to explicity designated

[8] Helson, *op. cit.*, this JOURNAL, 60, 1947, 1-30.

categories. Should these be fewer in number than required by the range of stimuli judged, the extreme categories would necessarily be indeterminate in magnitude. The integer serving as the interval-score in such cases would not represent the midpoint, but would lie closer to the proximal limit — the more restricted the scale-range, the greater this adjacency to the proximal limit, and consequently the smaller the slope of the psychophysical function in its extreme segments. Should instructions provide, however, that *S* has a single reference at the scale's neutral point and an indeterminate number of equally sized categories from which to devise his own optimum range, then the assumption concerning interval-size would not be violated, nor the end-effect occur. Furthermore, the problem of what shall constitute an adequate number of intervals would not present itself.

With these considerations in mind, the following instructions were devised for the remaining groups:

You will be presented with one weight at a time. After lifting it you are to tell me how heavy or light you think it is by giving it a scale-rating. The scale typed on the card in front of you is provided to help you in judging. [The typed scale had a neutral point at 500, and was structured in 100-step units up to +1600 (very, very, very, very, very, very, very, very, very, heavy) and down to —500 (very, very, very, very, very, very, very, very, very light). Beyond these points arrows, with heads in a distal direction, were drawn and prominently labeled, TO INFINITY.] Whatever you regard to be medium, that is, neither heavy nor light but halfway between heavy and light, call 500. Assign numerical scores to the various weights according to how far above or below this point you think they fall.

There are two things to keep in mind. First, this scale is not absolute, but only provided to guide you. Therefore, you may pick the score that you feel best expresses the heaviness or lightness of the weight. Express this to the nearest hundred. Secondly, the numbers on this scale are not the same as the physical weight of the objects you are lifting, since this is a scale for expressing your perception *of heaviness or lightness.* As such only you can decide where its upper and lower limits are. Therefore you need not hesitate in giving as your judgment whatever value you feel appropriate. The highest and lowest numbers we have included are +1600 and —500. *These are not intended to indicate the limits of the scale.* If a weight feels like 1,800 or 6,800, or 18,000 or —800 or —8,000, etc. do not hesitate to say so. Are there any questions?

Derivation of scores for analysis. Since several sets of instructions were used it is necessary to make explicit the manner in which the various ratings were transformed into a set of scores with a common reference. With both the Helson and expanded Helson rating-scales, the middle-most or neutral category was arbitrarily assigned a score of 10. Each verbal label above and below this was then assigned an appropriate value, each category differing from that adjacent to it by 1. Thus, for example, very heavy was assigned a value of 13, very light, 7. With the groups receiving the final revision of instructions, scores in the interval 450-549 were assigned a value of 10, each succeeding interval of 100 above and below this being designated as one category. Thus ratings varying between 850 and 949 were assigned a score of 14, those between —50 and +49, a score of 5.

Median heaviness scores were calculated for each weight for each subject.

Results and Discussion

Position of the indifference point. Table I presents the mean for 10 *Ss* of the median heaviness judgments for each weight. A score of 10, as

noted above, represents a judgment of medium, *i.e.* neither light nor heavy, but half way between light and heavy. Scores greater and less represent judgments of heavy and light respectively. The right-hand column presents for each group the stimulus-weight in grams corresponding to a judgment of 10.

Inspection of this table indicates that the predicted indifference point was approximated in Groups 1, 2, and 3. In none of the groups did the mean of the indifference points for the individual *S*s differ reliably from 280 gms. For Groups 4 and 6 it is significantly greater than the expected

TABLE I

MEAN HEAVINESS RATINGS OF THE SEVERAL STIMULUS WEIGHTS AND INDIFFERENCE POINTS FOR EACH OF THE EXPERIMENTAL GROUPS
(A score of 10 represents a judgment of medium)

Groups	Weight									Indifference point
	220	260	300	312.1	314.4	340	1010	1231	1950	
1	7.98	9.21	10.62	—	—	11.41	13.76	—	—	282.4
2	8.06	9.22	10.43	—	—	11.56	16.49	—	—	285.6
3	8.51	9.16	10.49	—	10.96	11.38	—	—	—	278.4
4	6.65	7.78	9.00	—	—	10.25	—	20.97	—	332.0
5	9.75	15.05	18.00	18.47	—	20.65	—	—	—	222.4
6	7.20	8.05	9.04	—	—	10.37	—	—	20.81	328.8

value ($t = 4.41$ and 2.34 respectively, $df = 9$, $p < .01$ and $< .05$), and for Group 5, significantly less ($t = 3.16$, $df = 9$, $p < .02$). The values for Groups 4 and 6 do not differ significantly from each other. It therefore appears that the principle of the adaptation-level holds only within certain limits. For the single condition described by the present curve of 'additional experience,' these would seem to be between 312.1 gm. presented 200 times at one extreme, and 1231 gms. presented once at the other. Further testing with intermediate values should yield a more precise statement of their position. Meanwhile, it can be confidently stated that all magnitude-presentation frequency-values between 314.1/100 and 1010/1.16 constitute a family of equivalent stimulus conditions.

The evidence that a limit exists suggests that Helson's mathematical expression should be revised. This is not feasible, however, until the apparent breakdown of the principle can be explained. The shift downward in Group 5 is probably atributable to a preponderance of heavy responses associated with a clear-cut negative time error in the records of 6 *S*s in this group. In two *S*s the effect was particularly marked. The judgments given began in the range 6 to 8 and progressed upward until, at the end of the session, they were in the high 40s and 70s respectively. Since a difference in physical magnitude existed between successive trials on only 12 occasions, formation of the necessary reference scale was undoubtedly a formidable task.

To a great extent S had to depend upon the immediately previous trial for his standard. Since on these trials the weights in most instances did not differ from those of the present trial, the influence of any time-error was maximized.

The upward shift in Groups 4 and 6 is not so readily accounted for. It appears that widely deviant stimuli have a value greater than those provided by the Helson formula. This is to say that they assume an anchoring function. It is not clear that S directs greater attention toward them, although it is plausible that, even with the random order of presentation, their obviously deviant properties may incline Ss to form rather specific expectations concerning their occurrence. Certainly either their extreme heaviness or their infrequent occurrence might enhance their emergence as figure.

Subsidiary considerations. While our primary concern has been to investigate the effect of varying the pattern of presumed-to-be equivalent experience upon the position of the indifference point, answers to certain other questions may be sought in the data. What effect might varying the pattern of experience have upon precision of judgment, the number of categories utilized in judging, and the slope and shape of the psychophysical function?

Precision. Certain differences in precision of judgment might be expected to occur. Precision, for example, should be greater for magnitudes farther removed from the indifference point.[9] Precision should also be lowest for Group 5, if our suggestion of a poorly articulated internal scale is valid.

TABLE II

INDEX OF PRECISION FOR WEIGHTS OF THE CORE, FOR EACH EXPERIMENTAL GROUP

Groups	Weights				Mean
	220	260	300	340	
1	5.1	5.5	5.8	5.5	5.48
2	6.5	7.1	7.1	6.2	6.72
3	3.4	3.0	3.0	3.2	3.15
4	7.6	10.0	8.6	9.2	8.85
5	15.2	11.8	16.2	12.2	13.85
6	8.6	8.8	9.1	8.5	8.75

Table II presents indices of precision for the various groups of the four weights of the common core experience. These indices consist of the mean number of categories applied in judging each weight; smaller scores indicate greater precision. Table III presents a summary of an analysis of variance of the data on precision. Since the scores in each group were independent of those in the remaining groups, it was necessary to use Block, Levine, and McNemar's modification of the conventional three-way analysis.[10] Examination of these tables indicates that, contrary to expecta-

[9] Volkman, Scales of judgment and their implications for social psychology, in J. H. Rohrer and Muzafer Sherif, (Ed.), *Social Psychology at the Crossroads*, 1951, 273-294.

[10] J. Block, L. Levine, and Quinn McNemar, Testing for the existence of psychometric patterns, *J. Abn. & Soc. Psychol.*, 46, 1951, 356-359.

tion, no significant differences in precision exist among the core weights. In contrast, a reliable difference does exist between the mean of the core weights (M_c) and that of the heavy weights (M_h)—1010, 1231, 1950)—in those groups making judgments with the latter: (1,24,6): $M_h - M_c = 1.89$; $t = 2.05$, $df = 39$, $p < .05$. The significant between-treatments (groups) F was further analyzed by a set of between-groups comparisons.[11] This last analysis indicates that only one of 15 differences (Group 5—Group 3) is reliable. Thus, while inspection of Table II suggests poor precision in Group 5, this impression is not statistically verified. A check on the raw data shows that the Group 5 scores are markedly inflated by the response of the two *S*s who display extremely large time-errors.

Extension of the judgmental scale. As stimulus-range is increased the number of categories utilized should increase or, if this number is fixed, the limits of the individual units should be extended. Since the instructions used were designed to maintain unit size relatively constant, an increase in the number of units was anticipated.

TABLE III

ANALYSIS OF VARIANCE OF PRECISENESS DATA

Source	Sum Sqs.	df	σ^2	F	p
Individuals	3,213.47	54	59.51	1.44	<.05
Weights	22.72	3	7.57	.18	>.05
Treatments	2,677.84	5	535.68	9.00	<.01
Interaction ($W \times T$)	156.24	15	10.42	.25	>.05
Residual	6,806.13	165	41.24		
Total	12,876.40	239			

It was not clear, however, that this prediction would hold for segments of the scale serving less than the total range. Accordingly, a simple analysis of variance was applied to an index of spread (the number of categories used in judging the core weights) and yielded significant differences among the several groups ($F = 4.62$, $df = 5/54$, $p < .01$). Paired comparisons, however, indicated these are not systematically related to total range.

Slope of the psychophysical function. Slopes were calculated for individual *S*s and a simple between-groups analysis performed which yielded a significant *F*-ratio ($F = 13.16$, $df = 5/54$, $p < .01$). Inspection of group means suggested this to be due to a difference between Group 5 and the remaining groups. This was confirmed by a re-analysis omitting Group 5 ($F = 0.48$, $df = 4/45$, $p > .05$). The steeper slope of Group 5 may reasonably be attributed to this Group's marked negative time error.

Shape of the psychological function. Final concern is directed toward the possibility that the conditions which effect a shift in the position of the indifference point also produce a change in the nature of the psychophysical function. This is of importance since it bears on the present conception of the judge as a logarithmic computer.

Straight lines of best fit were derived for the various groups by the Method of Successive Differences. These along with the formulae expressing them are given for all but Group 1 in Fig. 2.

[11] E. F. Lindquist, *Design and Analysis of Experiments,* 1953, 90-96.

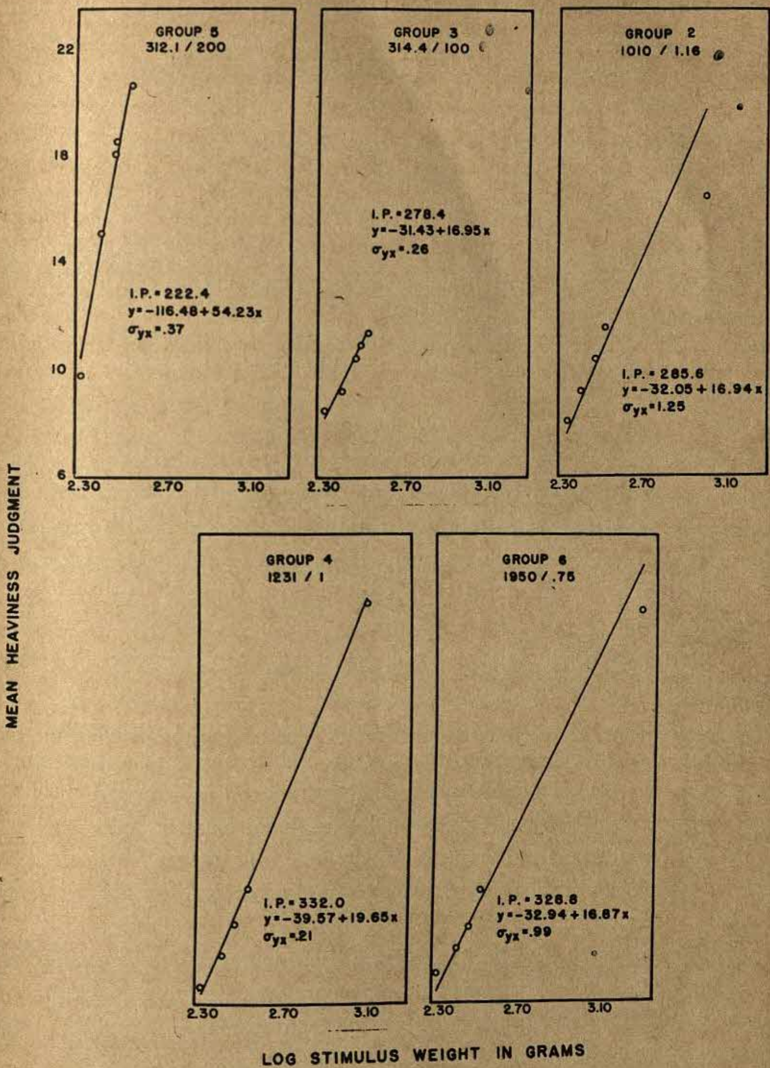


FIG. 2. JUDGED HEAVINESS OF STIMULUS WEIGHTS FOR FIVE OF THE EXPERIMENTAL GROUPS

A score of 10 represents medium. The straight lines are curves of best fit.

In the first groups tested, the fit was found to be good for Group 3, only one of the empirical values deviating significantly from the curve. The fit, however, was poor for Group 1, with the obtained values for 300, 340 and 1010 gm. all departing from the straight line by more than the standard error.

These results could be given at least two plausible interpretations: (a) The expected rectilinear relationship between frequency and stimulus-magnitude might

obtain only within certain limits, exceeded in Group 1; or (b) the rather rigidly structured instructions had resulted in a subjective scale too narrow to accommodate the range of magnitudes in Group 1. The mean number of categories used by Group 1 was 7.5 and the largest number was 9, although it was presumed that the instructions had provided for more categories if they were needed. To test the second interpretation, *i.e.* (b), additional categories were formally incorporated into the instructions and the remaining conditions of Group 1 repeated with Group 2.

The data of Group 2 more closely approximated a straight line than that of Group 1. Only one empirical value, that for 1010 gm., deviated from the curve by more than one standard error of estimate. The mean number of categories used was 9.1, with 3 *Ss* using 10 or more. The importance of proper instructions for the results of judging thus appears to be demonstrated.

Instructions were again modified toward greater flexibility and the remaining groups tested. All these yielded excellent rectilinear fits as inspection of Fig. 2 makes clear.

SUMMARY

Recent experimental work has shown that psychophysical judgments depend upon the existence of an internal scale ordered out of the judge's continuing experience with relevant stimuli. The present study approached the problem of relevance by exploring the effect of varying the patterns of 'equivalent experience' upon the predicted position of such a scale's indifference point.

Six groups of 10 *Ss* each made somewhat over 200 judgments of single stimuli with five weights. Four weights, 220, 260, 300, and 340 gm., each weight presented three times, constituted a 'core of experience' common to all groups. A fifth weight, different for each group and varied in the frequency of its presentation such that the predicted indifference point for every group was the same, constituted 'additional experience.'

Prediction of the indifference-point by Helson's formula proved to be correct when 'additional experience' was maintained within particular limits of stimulus-magnitude and frequency of presentation. With the particular core used, the limits were, on the one hand, between 312.1 gm. presented 200 times for each time the core was presented, and, on the other hand, 1231 gm. presented once for each core. A lighter fifth weight presented more frequently and heavier fifth weight presented less frequently produced a shift in the position of the indifference point downward and upward respectively.

A strong negative time-error, attributed to limited variability of experience, is associated with a significantly steeper slope of the psychophysical function for one group.

The shape of the psychophysical relationship was rectilinear for all groups, except one. In this case the non-rectilinear shape was shown to be a function of the instructions used.

THE EFFECTS OF ERROR-MAGNIFICATION AND MARKER-SIZE ON BIDIMENSIONAL COMPENSATORY TRACKING

By W. F. Battig, E. H. Nagel, and W. J. Brogden, University of Wisconsin

In compensatory tracking, the display presents only the error in tracking, the difference between the position of the control and the position of the target. If a cathode-ray oscilloscope (*CRO*) is used, there are two markers on the display screen: one is a fixed reference-mark, usually at the center of the screen, that represents zero error; the other is a movable marker, the location of which, relative to the reference-mark, represents the magnitude of the tracking-error. The task of *S* is so to manipulate the control that the movable and reference-markers coincide—that the signals from the control compensate the signals from the target.

The relative displacement of the error-marker may be altered in at least three ways: by changing the control-display sensitivity, by changing the target-amplitude, or by magnification of the tracking error. Control-display sensitivity is the relation between extent of displacement of the control and extent of displacement of the movable error-marker. Target-amplitude is the relation between extent of movement of the target and extent of displacement of the error-marker. Error-magnification represents the relation between magnitude of the error at the differential for the control and the target, and the magnitude of the error in terms of displacement of the moveable error-marker on the display. Error-magnification may be produced also by simultaneous and uniform increments in control-display sensitivity and target-amplitude. It is by this latter method that error-magnification was studied in the present experiment.

Both error-magnification and control-display sensitivity have been studied by Helson.[1] He reported differences in unidimensional compensatory tracking between two diameters of hand-wheel control (control-display sensitivity) and between two degrees of magnification of the field of the visual display (error-magnification). Swartz, Norris, and Spragg studied bidimensional pursuit-tracking as a function of crank-radius (control-display sensitivity) and found an optimal radius above and below which tracking-

* Accepted for publication October 25, 1954. Supported in part by a contract with the United States Air Force, monitored by the Psychology Branch, Aero-Medical Laboratory, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio.

[1] Harry Helson, Design of equipment and optimal human operation, this JOURNAL, 62, 1949, 473-497.

error increases.[2] The function obtained by Swartz, Norris, and Spragg is representative of the functions that might be expected to express the relation of each one of the variables of control-display sensitivity, target-amplitude, and error-magnification to magnitude of error in compensatory tracking. With the limitations imposed by acuity of vision and muscular adjustment, tracking-error might be expected to be large when the values of these variables are relatively small or relatively large, and at a minimum for some intermediate value. It is this hypothesis of the relation between error-magnification and magnitude of error in bidimensional compensatory tracking that the present experiment was designed to test. The experimental design provides also for investigation of the effects of marker-size, practice, and any interactions of the three experimental variables upon magnitude of tracking-error.

Method and Procedure

Apparatus. The apparatus has two stations separated by a partition, making it possible for two *S*s to perform the tracking-task simultaneously. The apparatus consists of (1) a target-course generator, (2) two *CRO* displays, (3) two control-sticks, (4) four electronic integrators for scoring tracking errors in azimuth and elevation for each of the two stations, (5) two clocks for scoring time-on-target in azimuth and elevation for each of the two stations, and (6) the necessary power-supplies and circuits for coupling of the components. The target-course generator is virtually the same as that used in the Dual Pursuit Apparatus.[3] A schematic block-diagram of the present apparatus is given in Fig. 1. A 1-r.p.m. synchronous motor rotates the metal cam. The two cam-followers, one for the azimuth selsyn, the other for the elevation selsyn, are separated by 180°. The angular position of the rotor of the transmitter selsyns is determined by the cam through the cam-follower. The signal of each transmitter selsyn is so modified by the position of the rotor of the appropriate differential selsyn on the azimuth or elevation axis of the control-stick that the difference in angular position of the two selsyns is represented as a deviation of the error-marker from the center of the *CRO* screen. The same signals serve as inputs to the integrators, and a separate integrated error-voltage is obtained for azimuth and elevation for each of the two stations. The reading on the integrator-meters is the summation of error over the 1-min. trial. The voltage reading is converted to deviation of the stick-position from the correct stick-position in degrees by use of appropriate calibration tables. The error-signals also operate relays that provide the power for the clutch of each electric clock-motor. The scoring area is variable and adjustment of controls for azimuth and elevation will determine the magnitude of the signals below and above which the clocks will and will not

[2] Paul Swartz, E. B. Norris, and S. D. S. Spragg, Performance on a following tracking task (modified *SAM* Two-Hand Coordination Test) as a function of radius of control cranks, *J. Psychol.*, 37, 1954, 163-171.

[3] W. F. Grether, A dual compensatory pursuit apparatus for use in psychological research. USAF Materiel Command, Dayton, Ohio, Technical Report No. 6036, 1950, 1-7.

operate. For the present experiment, the scoring area was set at 0.25 in. in each direction from the reference-mark for both azimuth and elevation.

A Variac transformer makes it possible to vary the voltage at the selsyns. Increase of the voltage results in a simultaneous and uniform increment in control-display sensitivity and target-amplitude, and thus results in magnification of the tracking-error. A voltmeter in this circuit provides for a direct reading of the voltage at the selsyns. Calibration of the voltage, in terms of displacement of the error-marker in inches on face of the scope for one degree of error between the selsyns on the

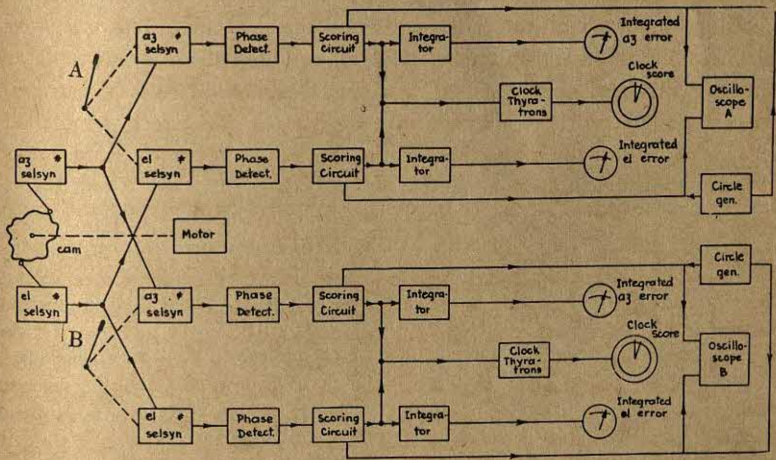


FIG. 1. SCHEMATIC BLOCK-DIAGRAM OF BIDIMENSIONAL COMPENSATORY TRACKING APPARATUS

A, B = control-sticks; az = azimuth; el = elevation; * = selsyn transmitters; ≠ = selsyn differentials used as resolvers.

control and the selsyns on the target-course generator, provides the means of varying error-magnification.

The target-course generator presents eight different courses without change of cam. These courses are provided by four starting positions and by the possibility of operating the motor in either clockwise or counterclockwise rotation. Only one target-course was used in the present experiment.

The *CRO*s are Dumont's type 304-*H*. A plywood panel painted black was placed over the front of each *CRO* to hide the controls. A rubber viewing hood, 9 in. long, was attached to each *CRO*. The end of the hood is molded to provide a good fit to the form of *S*'s head. The lucite overlay has at its center a 0.1-in. hole filled with luminous paint. Edge lighting at the top and bottom of the overlay provides sufficient luminance for the center dot to fluoresce and make this reference-marker approximately equal in brightness to the movable error-marker. The error-marker is a circle whose diameter was varied in the experiment by appropriate controls.

The joy-sticks have bicycle-type, rubber handgrips and are relatively lightly-loaded. The effective radius of rotation of the joy-stick when gripped with the hand is 7 in. Counterweights match the weight of the azimuth selsyn. None is needed for the

elevation selsyn since it is mounted on the frame. The joy-sticks may be adjusted in height by moving the ring-mount up or down a shaft mounted in a heavy cast-iron base. Screw adjustments of the feet of the base make it possible to keep the azimuth and elevation selsyns adjusted to zero when the stick is in its vertical position of rest. The height of the control-stick was held constant, but *S* was seated in a chair that could be adjusted in height according to his stature.

The apparatus provides for selection of all possible arrangements of stimulus-response compatibility between direction of movement of the control-stick and direction of movement of the error-marker. Direct stimulus-response compatibility was a condition of the present experiment. Movement of the control-stick to the right produced movement of the error-marker to the right; stick left, error-marker to left; stick forward, error-marker down; stick back, error-marker up.

The circuit for the electronic integrators is a modification of the type developed by G. A. Philbrick Researches, Inc. for their Analog Computer Systems. With appropriate adjustment and calibration the integrators have zero drift. Two rates of integration and four levels of meter-sensitivity are available. *E* selected these settings in terms of the performance of each individual *S* and the stage of practice.

Subjects. The 120 men, all students at the University of Wisconsin, served as *Ss*. They were paid for participation in the experiment.

Procedure. The six values of error-magnification used in the experiment were 0.08, 0.16, 0.24, 0.32, 0.40 and 0.48 in. of displacement of the error-marker from the reference-mark on the *CRO* screen for an error representing one degree of difference in rotation between the control-stick and the follower on the course-cam. The five diameters of the circular error-marker were 0.2, 0.5, 0.8, 1.1, and 1.4 in. There were four replications of each cell of the 6×5 factorial design. *Ss* were assigned to a particular cell of the design at random and the attempt was made to test two *Ss* simultaneously for each set of conditions, one at Station *A* and the other at Station *B*. Because of difficulties in making appointments and having *Ss* keep the appointments, it was not possible to meet this condition entirely. Of the total *N* of 120, 78 were tested in pairs and 42 were tested singly. Half the *Ss*, however, were tested on Station *A* and the other half were tested on Station *B*.

Verbatim instructions were given to each *S*. The task was explained and *S* was urged to do his best to maintain the position of the movable marker-circle that its center was coincident with the white dot at the center of the *CRO* display. The relations between direction of stick movement and direction of error-marker movement were given, and information about number of trials and rest-periods also was provided. Questions were answered if they involved clarification of the instructions. The chair-height was adjusted for each *S*, and *E* made certain that each *S* was in the proper position.

Fifteen 1-min. trials were given. There was a 20-sec. interval after each trial, except that between Trials 5 and 6 and Trials 10 and 11 the interval was 2 min. *E* signaled *S* vocally before he pressed the master switch to begin each trial. At the end of the trial *E* recorded the readings of the integrators and the clocks. The clocks had to be reset to zero by *E* before the next trial, but the integrator-meters were reset automatically by pressing the master switch to begin the next trial. *E* started the next trial when the appropriate intertrial interval had elapsed.

The voltage readings taken from the integrators must be transformed to tracking-

error in terms of the difference between actual stick-position and correct stick-position in degrees. The tracking-error for different values of error-magnification must be converted to a common reference since error-magnification also alters the magnitude of the signals to the integrators. Initial corrections were made for rate of integration and meter-sensitivity. Error-magnification of 0.20 in. per degree of tracking-error was taken as the reference. The adjusted raw score from the integrator was divided by two quantities. One of these is the ratio of the error-magnification at which the reading was made to the error-magnification of 0.20 in. per degree. The second quantity is the integrator reading in volts obtained for one of degree of tracking error over the 1-min. trial at an error-magnification of 0.20 in. per degree. These computations were made for each raw integrator score in azimuth and elevation for each trial for each *S*.

RESULTS

The primary measures of tracking are the integrated error-scores. Since the secondary measures of time-on-target in azimuth and elevation obtained

TABLE I

RESULTS OF ANALYSES OF VARIANCE

Source	df.	Azimuth Sum of squares	Azimuth Mean square	Azimuth F	Elevation Sum of squares	Elevation Mean square	Elevation F	Value of F at 5-% level
Total between-Ss	119	1341.73	—	—	982.32	—	—	—
Error-magnification (*EM*)	5	319.30	63.86	8.56	350.25	70.05	15.64	2.37
Marker-size(*MS*)	4	35.88	8.97	1.20	5.48	1.37	0.31	2.52
Apparatus-unit (*AU*)	1	1.15	1.15	0.15	21.04	21.04	4.70	4.00
EM×*MS*	20	261.60	13.08	1.75	146.00	8.30	1.85	1.75
EM×*AU*	5	10.80	2.16	0.29	31.35	6.27	1.40	2.37
MS×*AU*	4	106.60	26.65	3.57	44.60	11.15	2.49	2.52
EM×*MS*×*AU*	20	158.80	7.94	1.06	114.80	5.74	1.28	1.75
Error	60	447.60	7.46	—	268.80	4.48	—	—
Total within-Ss	1680	1772.82	—	—	1701.14	—	—	—
Trials (*T*)	14	530.60	37.90	59.22	465.08	33.22	53.58	1.69
T×*EM*	70	226.80	3.24	5.06	238.70	3.41	5.50	1.35
T×*MS*	56	26.88	0.48	0.75	22.40	0.40	0.65	1.35
T×*AU*	14	11.90	0.85	1.33	28.84	2.06	3.32	1.69
Pooled error	1526	976.64	0.64	—	946.12	0.62	—	—
Grand total	1799	3114.55	—	—	2683.46	—	—	—

from the clocks gave results similar to those of the integrated error-scores, only the latter results are reported. Table I presents a summary of the results of separate analyses of variance of the data for azimuth and elevation. These analyses follow the procedure outlined by Edwards.[4] For both the analyses of variance and subsequent statistical tests, the criterion of significance is the 5-% level of confidence.

The mean squares for the triple interactions of the sources for the within-*S*s sum of squares were computed and tested against the residual error. Since all of these interactions were far below the 5-% level of confidence, they were pooled with the residual error to provide a single

[4] A. L. Edwards, *Experimental Design in Psychological Research,* 1950.

error term. The results of the analyses for azimuth and elevation are quite similar. Of the main variables, error-magnification and trials (practice) have a significant effect upon tracking proficiency, but marker-size does not. Apparatus-unit is significant for elevation, although not for azimuth, which reflects a mismatch of the two units in the elevation component. Unit *A* consistently provides smaller measures of tracking-error than does Unit *B*. The interactions of apparatus-unit with marker-size and with trials are further reflections of difference between the two stations of the apparatus. Examination of the group-means of tracking-error for marker-size shows a non-systematic difference for the two units. Since there are no significant differences in tracking-error as a function of marker-size, the interaction of marker-size and apparatus-unit is not meaningful. In the case of the significant interaction of apparatus-unit and trials for elevation tracking-error, the effect is similar to that of apparatus-unit and elevation tracking-error. Apparatus Unit *B* gives consistently larger measures of tracking error than does Unit *A*. The difference between the two units is greatest early in practice and declines in magnitude as practice progresses. At the end of practice, Unit *B* still gives error-scores of greater magnitude than does Unit *A*. These results raise a question about pooling the data for the two units. Since there is no significant interaction of error-magnification and apparatus-unit, the relation between tracking-error and error-magnification is not materially affected by a difference between the two stations. Therefore, no breakdown by apparatus-unit of the relation between error-magnification and tracking-error is made.

There is, however, a significant interaction of error-magnification both with marker-size and with trials. These interactions indicate that the function of error-magnification and tracking-error varies as a function of marker-size and of trials. Further analysis of these effects, therefore, is necessary. Examination of the group-means for azimuth and elevation tracking-error as a function of error-magnification for each of the five marker-sizes shows no systematic effect of marker-size. The effect of this variable on the relation of error-magnification and tracking proficiency appears to be haphazard and random. No further presentation of these effects is given, therefore. The effect of practice, on the other hand, is systematic. Learning curves plotted for each condition of error-magnification, for both azimuth and elevation tracking-error, shows a reduction in tracking-error with practice. The amount of learning varies as a function of error-magnification. It is low for error-magnifications of 0.16, 0.24 and 0.32 in. per degree. It increases progressively with error-magnifications of 0.08, 0.40, and 0.48 in. per degree. The differences in tracking-error for the different magnitudes

of error-magnification are fairly consistent at all stages of practice. There is some crossing of the curves, particularly of those for 0.08, 0.24, and 0.32 in. per degree. The magnitude of the differences between the curves decreases as practice increases. The major gains with practice are complete by the tenth trial, and the curves are virtually parallel with each other and with the abscissa for the last five practice trials.

The effect of practice and error-magnification on azimuth and elevation tracking-error is shown in Fig. 2 where the group-means for each error-

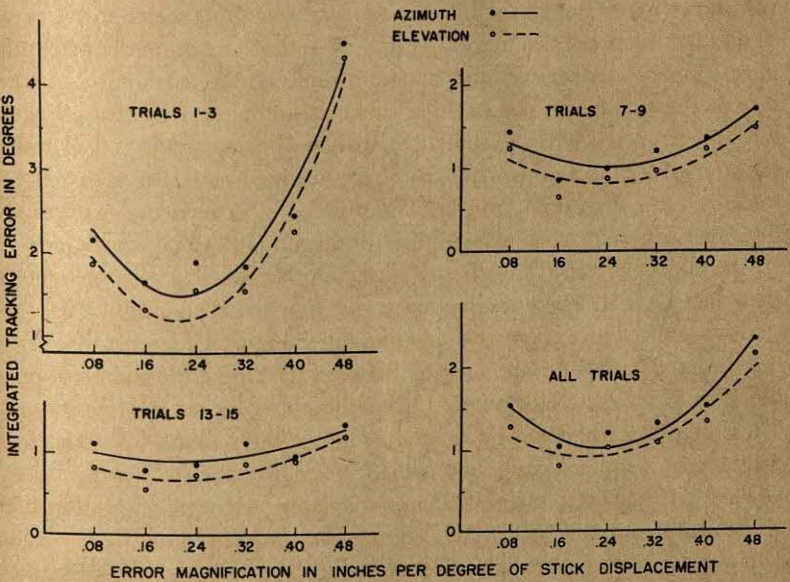


FIG. 2. THE RELATION BETWEEN TRACKING-ERROR AND ERROR-MAGNIFICATION WITH PRACTICE-LEVEL AS A PARAMETER

magnification are plotted for Trials 1–3, 7–9, 13–15, and all 15 trials. The smooth curves were obtained by a special method for fitting the polynomial, $y = a + bx + cx^2 + dx^3 \ldots$, to each set of data.[5] This method makes possible the successive fitting of terms of higher and higher degree, the testing for significance at each stage, and an analysis of variance. The fifth order was the terminal stage of the method in each case. The results of the analyses of variance indicate significant differences in tracking-error as a function of error-magnification except for the azimuth error-data for

[5] R. A. Fisher, and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research,* 4th ed., Hafner, 1953, 1-126.

trials 13-15. Significant first (linear) and second (quadratic) order components were found for each set of data except that of azimuth error for trials 13-15. For the latter set of data, only the quadratic component was significant. There were no significant contributions by the third, fourth, or fifth components for any of the data. Therefore, the curves of Fig. 2 are the fitted regression equations of the second degree orthogonal polynomial. These equations are given in parabolic form in Table II. The Chi-square test of goodness of fit yielded a Chi-square for each curve that is not

TABLE II

FITTED PARABOLIC REGRESSION EQUATIONS

Trials	Azimuth tracking error	Elevation tracking error
1-3	$y=3.43-1.42x+0.26x^2$	$y=3.08-1.40x+0.26x^2$
7-9	$y=1.72-0.48x+0.08x^2$	$y=1.53-0.48x+0.08x^2$
13-15	$y=1.01-0.16x+0.03x^2$	$y=1.01-0.27x+0.05x^2$
1-15	$y=2.23-0.82x+0.14x^2$	$y=1.51-0.40x+0.09x^2$

significant at the 99-% level of probability. The relation between tracking-error and error-magnification, therefore, is given adequate mathematical expression by the second-degree polynomial or parabola.

The parabolic form of the relation between tracking-error and error-magnification shows tracking-error to be at a minimum between values of error-magnification of 0.16 and 0.24 in. per degree for every condition of practice. Tracking-error increases in magnitude for error-magnifications of greater and lesser magnitude, and is greater for the larger error-magnifications than for the smaller ones. The results are very similar for azimuth and elevation tracking-error except that there is a consistent superiority of performance in elevation over performance in azimuth. The effect of practice is to reduce the magnitude of tracking error. The effect is differential, however, in that greater gains in performance are made at the large and small error-magnifications than for the optimal magnification between 0.16 and 0.24 in. per degree.

DISCUSSION

Although the diameter of the marker-circle is not a significant factor in determining magnitude of tracking-error, note should be taken of the display of the present experiment. A display consisting of a movable circular marker and a small reference-dot is not usual for compensatory tracking. The more ordinary display provides a small movable dot-marker that is adjusted in position to the center of a fixed reference-circle. Since the movement in both cases is relative to the two markers, there is no reason

to expect any variation in magnitude of tracking error as a function of diameter of a fixed reference-circle when the movable marker is a dot.

The finding that the relation between tracking-error and error-magnification may be expressed as a parabolic function at all stages of practice in the present experiment lends support to the general hypothesis concerning the relation between tracking-error and each of the variables of control-display sensitivity, error-magnification, and target-amplitude proposed in the introduction of this paper. It is likely that the relation between tracking-error and crank-radius (control-display sensitivity) found by Swartz, Norris, and Spragg could be adequate mathematical expression by a parabolic regression equation.[6] Other studies, particularly one involving target-amplitude as the independent variable, will be required to test the generality of the hypothesis.

The effect of practice upon the parabolic relation between tracking-error and error-magnification raises the question of whether this relation might not become linear, parallel to the abscissa, with an additional amount of practice. The 15 trials of the present experiment represent a relatively small amount of practice. Bidimensional compensatory tracking is a task of some difficulty, and further reduction in tracking-error is a distinct possibility. The fact that all learning in the present experiment was completed by the tenth trial is an argument against further improvement with practice. The stable performance during trials 11-15 may, however, represent a plateau rather than a terminal asymptote of the learning curves. Many of the *Ss* of the present experiment reported fatigue toward the end of the practice-period, which suggests that performance might continue to improve with additional practice distributed more widely over time. The primary question, however, is whether the differential effect of error-magnification upon tracking performance will persist or be eliminated with extensive practice. This question is worthy of experimental investigation.

The consistently larger tracking-error in azimuth over that in elevation must be due to a relatively greater ease in making positioning movements of the hand and arm in the forward-backward dimension than in the left-right dimension. Other investigators have found either a similar difference or no difference at all. Grether reported superior performance in a continuous pursuit-task for fore-and-aft movements of a stick-type control, and of a wheel, over side-to-side or rotary movements of the same controls.[7] Corrigan

---

[6] Swartz, Norris, and Spragg, *op. cit.*, 163 f.

[7] W. F. Grether, Efficiency of several types of control movements in the performance of a simple compensatory pursuit task. In P. M. Fitts (Ed.) *Psychological Research on Equipment Design*, U. S. Government Printing Office, 1947, 1-276.

and Brogden found no difference between the precision of linear pursuit-movments in the planes perpendicular to (0°) and parallel to (90°) the frontal surface of the body.[8] Helson and Howe obtained equivalent measures of tracking for a handwheel-control when it was rotated in the vertical, horizontal, and oblique positions.[9] Whether differences in precision of tracking performance occur between movements in the vertical or horizontal appears to be a function of the specific nature of the control. From the results of the present study and those of Grether, it appears that with a stick-type control, fore-and-aft movements produce performance superior to that produced by side-to-side movements.

SUMMARY

The effects of six values of error-magnification (0.08, 0.16, 0.24, 0.32, 0.40, and 0.48 in. per degree) and five values of moveable marker-diameter (0.2, 0.5, 0.8, 1.1. and 1.4 in.) on bidemensional compensatory tracking were studied. The four *S*s in each cell of this 6 x 5 factorial design were given 15 2-min. tracking trials. Analysis of variance shows no effect of movable marker-diameter upon magnitude of tracking error. Error-magnification and practice are significant sources of variation and there is a significant interaction of these two variables, the effect of practice being to reduce tracking-error for each condition of error-magnification. The amount of learning, however, varies as a function of error-magnification. In all cases, the effect of practice is completed by the tenth trial. An orthogonal polynomial of the second degree fits the relation between tracking-error and error-magnification for the stages of practice represented by Trials 1–3, 7–9, 13–15, and all trials. The parabolic regression equations show optimal performance at error-magnification between 0.16 and 0.24 in. per degree, with tracking-error increasing in magnitude at lesser and greater values of error-magnification. The effect of practice is a reduction of the differences in tracking-error as a function of error-magnification. The results for azimuth and elevation tracking-error are the same except for a consistent superiority of performance in the elevation component over that in the azimuth component. The results are discussed in relation to relevant literature and in terms of a general hypothesis concerning a parabolic relation between tracking-error and the variables of control-display sensitivity, target-amplitude, and error-magnification.

[8] R. E. Corrigan, and W. J. Brogden, The trigonometric relationship of precision and angle of linear pursuit-movements, this JOURNAL, 62, 1949, 90-98.

[9] Harry Helson, and W. H. Howe, A study of factors determining accuracy of tracking by means of handwheel control, OSRD Report No. 3451, The Foxboro Co., 1943 (PB40617), 1-45.

MEMORY FOR ITEMS IN A MATRIX

By SCARVIA B. ANDERSON, Naval Research Laboratory, and*
SHERMAN ROSS, University of Maryland

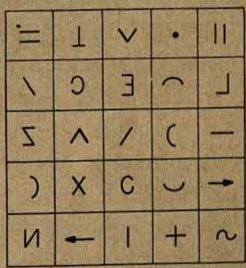
This study is concerned with a five by five square matrix. Let us consider for a moment some of its properties. It consists of 25 cells, each of which is uniquely specifiable in terms of its location with respect to every other cell. Taking an analogy from symbolic logic, each cell in the matrix may be considered a logical variable. A logical variable is in the strictest sense a place-holder into which the names of various values may be inserted. For example, taking the cell in the upper left-hand corner of this matrix, we can imagine that it is filled with a letter, a word, or any number of things (see Fig. 1). The same is true of every other cell in the

x	e	a	g	t
r	l	i	s	b
u	o	c	z	h
d	j	m	q	n
w	f	k	p	y

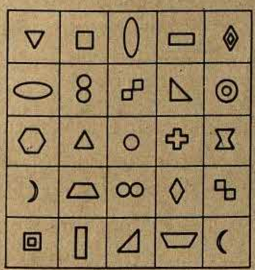
LETTERS

85	69	27	31	23
17	41	58	74	35
75	26	61	82	37
78	62	79	24	19
86	48	54	49	38

NUMBERS



SYMBOLS



GEOMETRIC FIGURES

get	top	ran	lot	bed
yes	big	hit	now	cat
fun	low	raw	bag	fix
lap	tin	ham	car	sad
new	fur	rip	pen	dog

WORDS

FIG. 1. EXAMPLES OF THE STIMULUS-MATERIALS

matrix. Regardless of the substitutions, any given variable remains fixed in location with respect to the other variables.

* Accepted for publication November 4, 1954. The authors are indebted to Dr. Omar K. Moore, Yale University, and to Dr. T. G. Andrews and Dr. R. C. Hackman, University of Maryland, for helpful suggestions and criticisms.

The present study, exploratory in nature, was designed to determine whether there are any relatively invariant relationships between the learning of material and the spatial location of the place-holders into which that material is inserted. When stimuli are presented simultaneously in the matrix indicated here, are there any favored locations in terms of correct responses? If so, can they be systematically ordered according to some simple principle? If there is such an order, how can we account for it? Bartlett, commenting on Binet's concept of *idée directrice*, has said that when material is presented, the subject (*S*) discovers what he takes to be its rule of arrangement.[1] Thereafter the rule itself becomes predominant and fashions his recall. We were interested in what rules of arrangement *S* attributes to a square matrix.

METHOD AND PROCEDURE

The experiment has its origins in studies on visual imagery reported by Woodworth and Hebb.[2] In these studies, *S* was presented with a 4x4 square matrix containing letters. He was told to study it until he had an image of the letters in the entire square. The hypothesis tested and unsupported was that if *S* really had an image of the square, it would not matter in what direction he was asked to report the letters. The method of stimulus-presentation used in these earlier experiments seemed adaptable for use in the problem area with which we are concerned.

Materials. The values of the 'variables' used in the present experiment were of five kinds: Letters, Numbers, Symbols, Geometric Figures, and Words. The Letters included all of the letters of the alphabet except 'v,' which was omitted because of its similarity to an item under Symbols. The 25 two-digit Numbers were chosen from a table of random numbers and no number ended in zero. The Symbols consisted of logical symbols and symbols of the type used in the Wechsler-Bellevue Digit Symbol Test. We believed that these represented the most unfamiliar material. The items classified as Geometric Figures were informally pre-tested for discriminability. The Words were all three-letter words beginning and ending with consonants. They had similar frequency counts according to the Thorndike-Lorge criterion.[3] None of the words rhymed, and the matrix ordering was carefully inspected to insure that adjacent words did not seem to have strong associative value. The 25 items on each kind of matrix were pre-tested to see that they could be reproduced in the time alloted for recall. The time was adequate.

Five forms for each kind of material were constructed by random assignment of items to the 25 squares of the matrix. No item appeared in the same square on any two forms. Fig. 1 shows one form of each kind of material. Response-cards were blank matrices.

Subjects. The *Ss* were five enlisted men from the U.S. Navy who had GCT scores between the 76th and 86th percentiles. They ranged in age from 17 to 21 yr., and their educational backgrounds were similar.

---

[1] F. C. Bartlett, *Remembering,* 1932, 52.

[2] R. S. Woodworth, *Experimental Psychology,* 1938, 42; D. O. Hebb, *The Organization of Behavior,* 1949, 36.

[3] E. L. Thorndike and Irving Lorge. *The Teacher's Word Book of 30,000 Words,* 1944.

Procedure. S was brought into the laboratory for five 1-hr. sessions. The sessions were separated by at least 5 hr. During each session S was given five trials with one form of each of the five kinds of material (a total of 25 trials). The order of presentation of materials and forms was randomized with the usual restrictions met in Latin Square design. S_1 during Session 1 might receive five trials on Form *a* of the Letter matrix, followed by five trials on Form *c* of the Geometric Figure matrix, and by five trials on Form *b* of the Symbol matrix, etc.

Instructions. At the beginning of the first session, the following instructions were read to S.

We wish to see how well you remember things. We are going to ask you to look at some symbols on a card and then try to recall them by writing them on another card. You will have five trials with each card. For example: If this is your stimulus-card, you will be allowed to look at it for 1 min. Then you will be told to turn it over and write on a report card what you remember. At the end of another minute, you will be told to turn over the card you are writing on. We'll repeat this procedure five times with each stimulus-card.

When I say 'Go,' turn over the stimulus-card and study it. Try to learn what is on it. As soon as I say 'Stop,' turn over the stimulus-card and try to fill in the report-card; when I say 'Stop' again, turn over the report-card. Are there any questions? [If so, they were answered.] 'Ready,' 'Go.'

One minute was allowed for each stimulus-presentation and 1 min. for filling out each response-card. Thirty seconds were allowed between trials with the same matrix and 2 min. between different matrices.

Results

Fig. 2 shows the total number of correct responses for all the Ss for all trials for the five kinds of material. Every figure represents a total of 3125 possible correct responses.

With Letters, Numbers, Geometric Figures, and Words there is a steady row-by-row decrease in correct responses, from the first through the fourth rows. The fifth row for Letters, Geometric Figures, and Words, however, appears to show about the same number of correct responses as the fourth row; and with Numbers the fifth row shows a few more correct responses than the fourth row. The Symbol matrix shows a somewhat different pattern of correct responses. Note the chasm at the center of the figure and the slight emphasis in terms of correct responses on the first two columns. With all five kinds of material, however, the first position in the first row shows a relatively large number of correct responses.

In terms of total number of correct responses for all Ss and all trials, the different kinds of material rank as follows: Letters, 2220 correct responses; Symbols, 1803; Words, 1747; Numbers, 1519; and Geometric Figures, 1432.

The learning curves for the various positions with the five kinds of material are shown in Figs. 3–7. They show which items were learned first, and the improvement in various items following initial recall is reflected

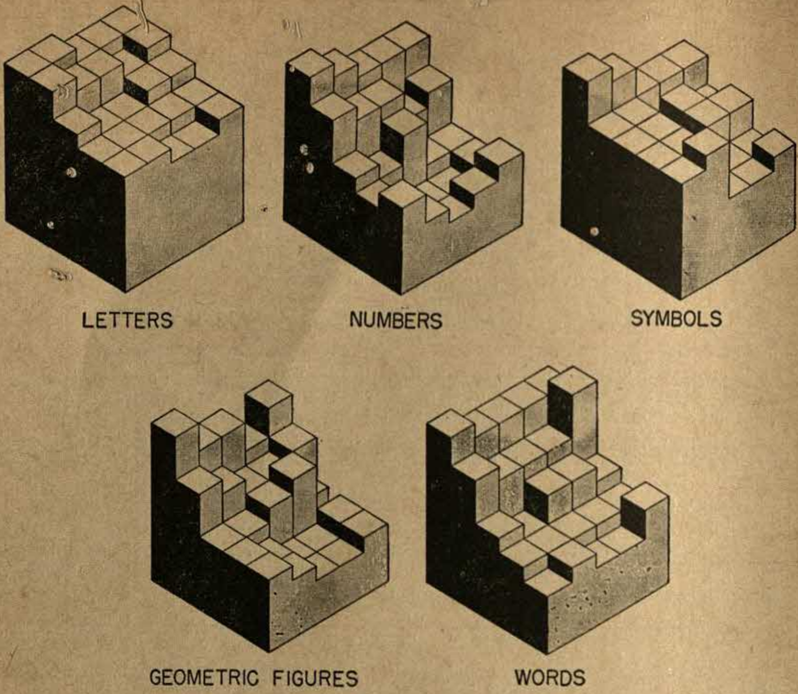


FIG. 2. CORRECT RESPONSES FOR THE VARIOUS TYPES OF MATERIALS

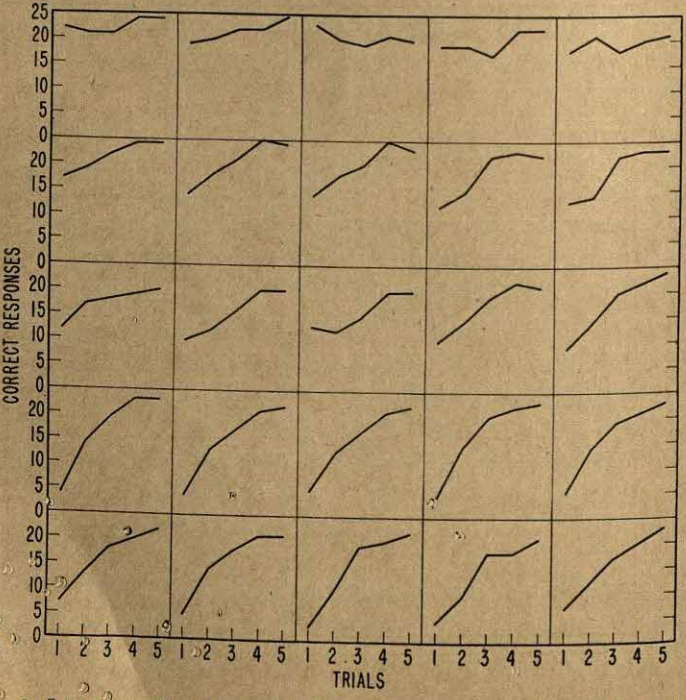


FIG. 3. LEARNING CURVES AT EVERY POSITION IN THE LETTER MATRICES

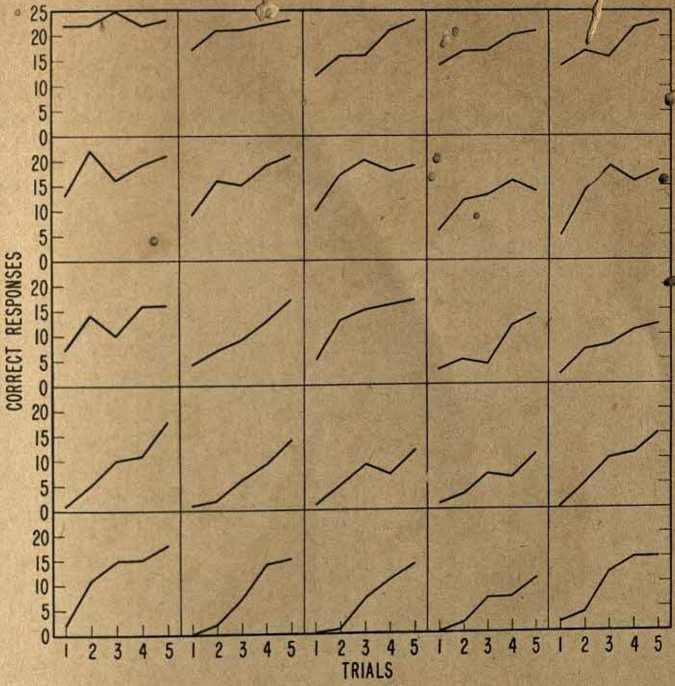


FIG. 4. LEARNING CURVES: NUMBER MATRICES

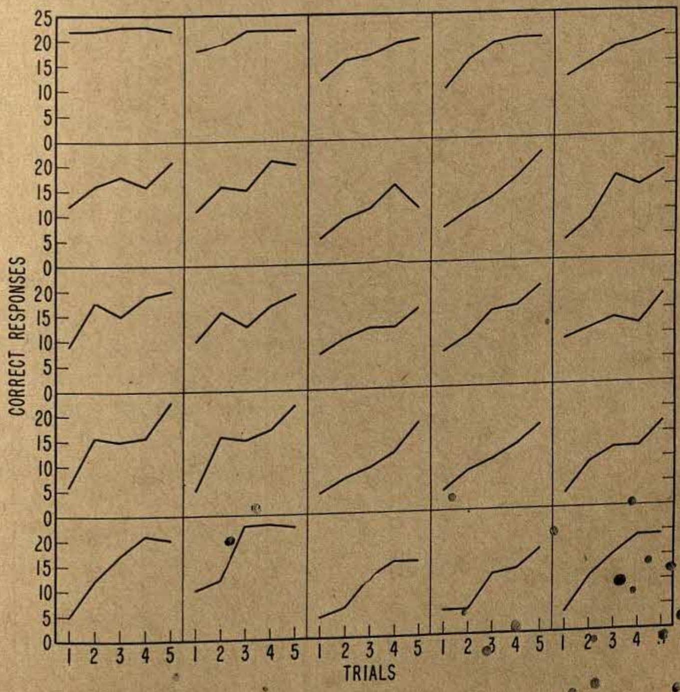


FIG. 5. LEARNING CURVES: SYMBOL MATRICES

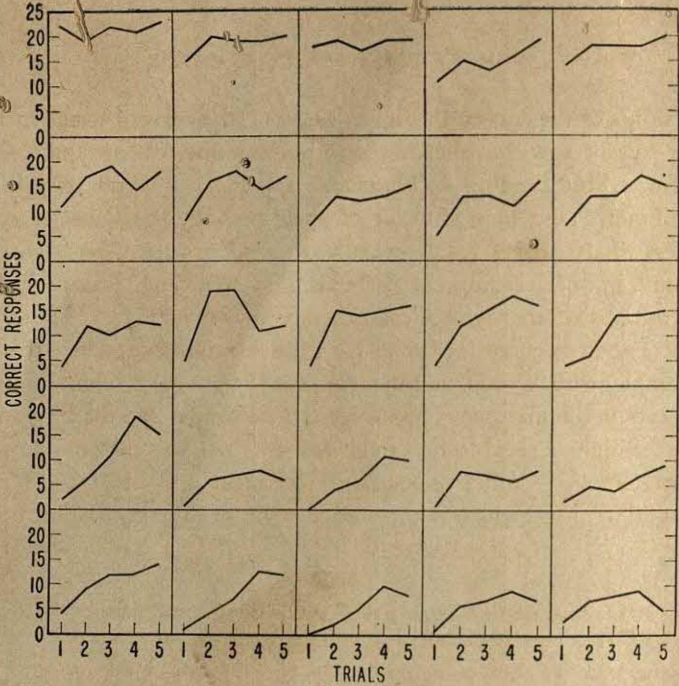


FIG. 6. LEARNING CURVES: GEOMETRIC FIGURE MATRICES

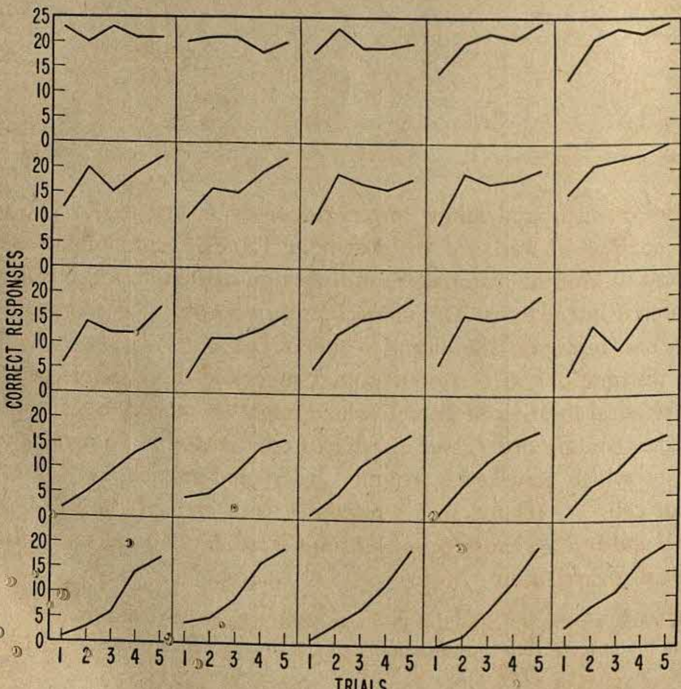


FIG. 7. LEARNING CURVES: WORD MATRICES

in the slopes of the curves. In Fig. 3 (Letters), the curves become markedly steeper row by row, but the curves within any one row are quite similar. The same is true for Fig. 4 (Numbers) and Fig. 7 (Words). With Geometric Figures, the most difficult of the materials, the same increase in steepness of the curves from row to row is not apparent, but the within-row similarity of learning curves holds (Fig. 6). The curves for the 25 item positions of the Symbol matrices are shown in Fig. 5. The row-by-row increase in steepness of curves is a little less pronounced here than it is, for example, with Words, and there appear to be some column differences. The curves in the first two columns seem to be similar and the curves in the last three columns seem to be similar row by row. This discontinuity effect between columns 2 and 3 seems fairly consistent.

Over-all improvement was analyzed in terms of difference scores (correct

TABLE I

Analysis of Variance in Terms of Improvement Scores

Correct responses on fifth trial minus correct responses on first trial.

Source of variation	Sum of squares	df	Mean square	F
Materials	37.91	4	9.48	6.08*
Ss	24.03	4	6.01	3.85*
Positions of Items	340.77	24	14.20	9.10*
M×S	68.27	16	4.27	2.74*
M×P	173.85	96	1.81	1.16
S×P	218.53	96	2.28	1.46*
M×S×P	599.97	384	1.56	
Total	1463.33	624		

* Significant at 0.01-level.

responses on fifth trial minus correct responses on first trial). The results of the analysis of variance are shown in Table I. Significant effects are attributed to kind of material, *S*, and position of item to *S* x position and *S* x material interactions. The interaction between material and position of item is not, however, significant in this analysis.

The distributions of correct responses suggested three factors pertaining to the physical location of stimuli which might be related to *S*'s responses: row order, column order, and number of adjacent cells. To investigate the extent to which row order, column order, and number of immediately adjacent cells are related to *S*'s responses with each of the five kinds of material and at different stages of learning, a rather unorthodox correlational analysis was carried out.

First, each of the 25 stimulus-positions was assigned a score from 1 to 5, in terms of the order (from top to bottom) of the row in which it appeared. Thus each

of the items in the first row was assigned a score of 1; each of the items in the second row, a score of 2; etc., as follows:

1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5

To study the relationship between learning and column order, the items in the first column on the left were assigned scores of 1, etc., as follows:

1	2	3	4	5
1	2	3	4	5
1	2	3	4	5
1	2	3	4	5
1	2	3	4	5

Every stimulus-position has three, five, or eight cells immediately surrounding it. Corner cells have three cells adjacent to them; other peripheral cells, five; central cells, eight. Scores of 3, 5, or 8 were assigned to every position in the matrix to test the hypothesis that intereference by contingent items might be a factor contributing to differential learning of items in the various positions in the matrix. The scores were as follows:

3	5	5	5	3
5	8	8	8	5
5	8	8	8	5
5	8	8	8	5
3	5	5	5	3

The arbitrarily assigned scores based on these three factors were correlated with number of correct responses in each position of the matrix for each kind of material for each trial. The resulting *r*s are shown in Table II. For all materials and almost every trial, row scores are significantly correlated with correct responses.[4] With Numbers and Geometric Figures the correlations between correct responses and row designations remain very high throughout the five trials. With Letters, Symbols, and Words, row scores show a decreasing relationship to response scores as the number of trials increases. When the stimuli are Symbols, row, column, and interference scores are related to subjects' response scores at one stage of learning or another.

Conclusions

From the patterns of correct responses and the correlations, we have noted some apparent consistencies of response regardless of kind of material; the analysis of variance of improvement scores fails to yield a significant material x position interaction. We therefore think it is safe to say

[4] The *t*-test was used for testing the significance of these small sample *r*s.

that with several different kinds of material, there are some fairly constant relationships between learning and spatial location.

With each kind of material 'upper left-hand corner' is a favored location. On the first two trials with each kind of material 'first row' is a favored location. Furthermore up to the point when almost complete learning has taken place, number of correct responses generally appears to be a decreasing function of order of row. There is one exception, Symbols.

The most prevalent pattern of correct responses seems to be directly related to the row reading habit—highly reinforced since childhood. *Ss*

TABLE II

Correlations of Correct Responses with Item Position and Number of Interfering Items

		Trial 1	2	3	4	5
Row (1–5)	Letters	−.93‡	−.86‡	−.42*	−.51†	−.15
	Numbers	−.88‡	−.87‡	−.72‡	−.79‡	−.76‡
	Symbols	−.68‡	−.56†	−.40	−.40	−.25
	Geom. Fig.	−.85‡	−.86‡	−.79‡	−.69‡	−.85‡
	Words	−.89‡	−.95‡	−.91‡	−.74‡	−.57†
Column (1–5)	Letters	−.11	−.20	−.03	−.16	.00
	Numbers	−.24	−.26	−.17	−.19	−.33
	Symbols	−.41*	−.58†	−.34	−.47*	−.36
	Geom. Fig.	−.17	−.20	−.28	−.23	−.22
	Words	−.11	.10	.15	.29	.31
No. interfering items (3, 5, 8)	Letters	−.21	−.23	−.08	.24	−.16
	Numbers	−.27	−.21	−.34	−.37	−.37
	Symbols	−.30	−.27	−.57†	−.41*	−.27
	Geom. Fig.	−.37	−.02	−.10	−.30	−.18
	Words	−.27	−.05	−.10	−.26	−.31

* Significant at 0.05 level. † Significant at 0.01 level. ‡ Significant at 0.001 level.

started 'memorizing' at the upper left-hand corner and learned the first row; then they shifted to the second row and started to learn it; and so on until the end of a trial. At the beginning of the next trial, they probably began again with the first row. With such a procedure, the amount of practice each row received would be a decreasing function of order of rows. We wonder what sort of distribution of correct responses we would get with *Ss* whose previous reading experience was only with Chinese orthography!

Correct responses to Symbol stimuli are distributed in a pattern somewhat different from that of correct responses to the other kinds of material. It seems reasonable to assume that many of the Symbols were unfamiliar to our *Ss*. Many of the items were hard to classify and hard to 'name.' Our *Ss* did not seem to adopt a uniform procedure for learning them. In ad-

dition, though no attempt has been made to dimensionalize the Symbol stimuli with respect to similarity, it does appear that there is more similarity between Symbol items than between items within any of the four other kinds of material. The similarity hypothesis is compatible with the fact that the Symbol matrix is the only one of the five kinds of material with which significant correlations were found between correct responses and number of surrounding items. Unfamiliarity, similarity, or both may have disrupted the more general correct response pattern.

Astin in a follow-up study using 9 *S*s and the Word and Symbol matrices obtained patterns of correct responses similar to ours and similar learning curves for the various cells.[5] He is able to supply an additional and interesting piece of information; namely, temporal order and accuracy of response are highly correlated (rho equals 0.90 for the Word matrices and 0.94 for the Symbol matrices).

Our experiment was exploratory, the number of *S*s was small, and our analysis of results was tentative. With the support of Astin's results, however, we believe that we are warranted in concluding that stimulus-items presented simultaneously in a 5 x 5 square matrix yield in learning some definite positional relationships—related perhaps to reading habits—which hold for various types and kinds of materials. We suggest for future study that the investigation of two variables, unfamiliarity of substitutions and similarity between substitutions, may resolve these relationships.

[5] A. W. Astin, Response ordering and recall of stimulus-items arranged in a square matrix, University of Maryland, unpublished manuscript, 1954.

EARLY LEARNING AND THE PERCEPTION OF SPACE

By JAMES DREVER, University of Edinburgh

Hebb has drawn a distinction between early and late learning.[1] During early learning, he suggests, organization occurs in the non-specialized cortical areas, and this organization acts as a basis for the perceptual skills and insights upon which later learning in part depends. Though of wide generality, this suggestion is not merely speculative in that it can be used to make predictions about subjects who have or have not had opportunity for certain kinds of early learning.

In stating his case Hebb says that the work of Senden[2] and Riesen[3] is fundamental to his argument. He also makes some use of his own work on rats, though this suggests that it is with the higher animals, presumably because of their greater proportion of unspecialized cortex, that the distinction between early and late learning is important.[4] Later work, however, confirms the existence of the distinction at the level of the rat.[5]

Now the evidence from the work of Senden and Riesen is not, as Hebb would be the first to admit, by any means conclusive. Senden gathered together reports on the behavior of patients operated on for congenital cataract; people, that is, who become able to see in late childhood or early adult life. But Dennis has already called attention to what he calls the "unquantitative as well as incidental nature of the observations" in most of these reports,[6] and Hebb himself tends to make use only of those cases whose subsequent visual learning was much impaired. Even in the use of this selected group he has been criticized by Wertheimer on the grounds that he exaggerates the frequency with which certain symptoms are reported by Senden.[7] The condition described as congenital cataract is itself very

* Accepted for publication September 14, 1954.

[1] D. O. Hebb, *The Organization of Behavior,* 1949, 109-120.

[2] M. von Senden, *Raum- und Gestaltauffassung bei operierten Blindgeboren vor und nach der Operation,* 1932, 135-141.

[3] A. H. Riesen, The development of visual perception in man and chimpanzee, *Science,* 106, 1947, 107-108.

[4] Hebb, The innate organization of visual activity: I, *J. Genet. Psychol.,* 51, 1937, 101-126; II, *J. Comp. Psychol.,* 24, 1937, 277-299.

[5] W. E. Bingham and W. J. Griffiths, Jr., The effect of different environments during infancy on adult behavior in the rat, *J. Comp. & Physiol. Psychol.,* 45, 1952, 307-312; R. H. Forgus, The effect of early perceptual learning on the behavioral organization of adult rats, *J. Comp. & Physiol. Psychol.,* 47, 1954, 331-336.

[6] Wayne Dennis, Congenital cataract, *J. Genet. Psychol.,* 44, 1934, 340-350.

[7] Michael Wertheimer, Hebb and Senden on the rôle of learning in perception, this JOURNAL, 64, 1951, 133-137.

variable and by no means involves total blindness in every case. Sensitivity to light is often present and some degree of object vision may also be possible. With these qualifications, however, there is still a certain amount of evidence, and a theory such as Hebb suggests may be necessary to explain the cases he describes.

Riesen's evidence from anthropoid apes reared in darkness is also relevant so far as it goes, but is at present weakened by doubt as to the amount of actual physical deterioration that may have taken place. Rearing in dim or intermittent light seems to be proving a more satisfactory technique.

Now it is much easier to find *S*s who have lost their sight than *S*s who have gained it, and evidence from a study of the blind may be relevant here, particularly evidence from a comparative study of *S*s who have lost their sight at different ages. A good deal of such evidence already exists. Schlaegel sums up earlier work on visual imagery in blind *S*s, and concludes from these and from his own studies that visual imagery, so far as it can be determined by introspection, is available to *S*s who have been able to see up to the age of about 6 yr., but not to any extent in cases where blindness has occurred earlier.[8]

Duncan using a finger-maze investigated the learning of blind and sighted *S*s.[9] She concluded "it seems that the present degree of vision is of much less importance, so far as maze learning ability is concerned, than the amount or duration of visual experience." The visual experience in question was of course early visual experience.

Worchel investigated the space-perception and orientation of (1) early blind, (2) late blind, and (3) sighted *S*s.[10] He found that the kinds of performance which he tested tended to improve from Group (1) through Group (2) to Group (3). This present investigation in part repeats and in part continues Worchel's work.

Subjects. Two groups of 37 *S*s, matched for age, sex, and intelligence, served in the study. One group was blind and the other sighted. The age-range of the *S*s selected was 12 yr. 0 mo. to 19 yr. 6 mo. A boy 9 yr. 8 mo. old was accidentally included in the blind group and gave results comparable to those from the older *S*s, hence a sighted boy of the same age was brought in to match him. The mean age of the blind group was 14 yr. 11 mo. and of the sighted 15 yr. 3 mo.

The mean Terman-Merrill *IQ* of the blind group was 111.7 with a range of 88 to 154; and of the sighted group, based on a Moray House group-test, 116.1 with a

[8] T. F. Schlaegel, Jr., The dominant method of imagery in blind as compared to sighted adolescents, *J. Genet. Psychol.*, 83, 1953, 265-277.

[9] B. K. Duncan, A comparative study of finger maze learning by blind and sighted subjects, *J. Genet. Psychol.*, 44, 1934, 69-95.

[10] Philip Worchel, Space perception and orientation in the blind, *Psychol. Monog.*, 65, 1951 (No. 332), 1-28.

range of 95 to 150. Taking their handicap into account, it seems reasonable to suppose that the blind group was at least as intelligent as the sighted despite their slight inferiority in average score.

In addition the blind *Ss* were divided into two matched groups. The early blind group, 19 in number, contained the *Ss* who were congenitally blind and those who, by the age of 4 yr. could not distinguish objects of any sort. Mere sensitivity to light was not regarded as seeing for the purposes of the investigation.

The late blind group, 18 in number, contained the remainder of the blind *Ss*. These *Ss* had lost their sight at various ages, but none had become blind within the 2-yr. period preceding the investigation. Three cases, although blind for most practical purposes, could still distinguish steep gradients of illumination and had thus some slight degree of object-vision.

Causes of blindness were various, and this gave rise to a complicating factor in all *Ss* not born blind; namely, the rate of onset of blindness.[11] In traumatic cases dating was easy, as it was with blindness following an acute infection, but in other cases the transition from seeing to not seeing occupied many years, and to establish the point at which an *S* became blind for the purposes of the investigation might well have been difficult. Fortunately full case-histories were available, and the headmaster of the school, who was keenly interested, gave invaluable help at every stage. In particular, he went over the case histories in detail, and the estimates for onset of blindness are his.

Since a part of the investigation which is not being reported here involved taking *EEG*s from all *Ss* throughout the experiment, it was thought wise to exclude those cases whose blindness was due to brain surgery or other central causes. In point of fact one such case is included where optic atrophy was not accompanied by *EEG* or other neurological abnormality.

The mean age of the early blind group was 14 yr. 11 mo. and of the late blind 14 yr. 10 mo. the mean *IQ* of the early blind group was 112.2 and of the late blind 111.05. Using these *Ss*, it is possible to compare the performance of the blind with that of the sighted, the performance of the early blind with that of the late blind, and the performance of either blind group separately with that of the sighted.

Problems. Every *S* was given three sets of problems in the tactile-kinesthetic modality—the sighted *Ss* having blindfolds over their eyes. These problems were: (1) Figure-recognition; (2) Orientation; and (3) Classification. In addition Hunter used some of the *Ss* in an investigation of the tactile-kinesthetic perception of straightness to which reference will be made later.

(1) FIGURE-RECOGNITION

(a) Method. Recognition of figures was a repetition of one of the experiments reported by Worchel. It was included for two reasons, one general the other particular. The general reason derives from the fact that we too often find in psychology

[11] They were diagnosed as follows: Glioma 5, congenital cataract 4, iridocyclitis 4, buphthalmia 3, keratitis 3, macular dystrophy 3, ophthalmia neonatorum 3, trauma 3, uveitis 3, choroido-retinitis 2, retinal detachment 2, retinitis pigmentosa 1, optic atrophy (post-operative) 1.

investigations which cover the same ground and seem to belong together but variations in design or technique make a strict comparison of results difficult. Providing an overlap makes it possible to treat this investigation and Worchel's earlier one as a single unit. The particular reason springs from the *EEG* part of the work. It was desirable to have a test whose visual properties were already known in case the new tests turned out not to have these properties.

The apparatus consisted of two sets of small wooden blocks ¼ in. in thickness. One set was used as the stimulus-object and the other for eliciting the response. The stimulus-set contained two parts each of a circle, semi-circle, ellipse, square, rectangle, trapezoid, and triangle. The response set consisted of these same figures presented complete. Each *S* was instructed as follows:

> I am going to give you two wooden blocks one in each hand. You can turn them around in your hands but you must not put them together. When you are ready I shall hand you four other blocks one at a time. You can handle these as you like. At the end I want you to tell me which of the four you could have made if you had placed the first two blocks together side by side.

S was then given the first two blocks. When he indicated that he was ready, generally within 20 sec., they were removed and the response-series begun.

(b) Results. The results for this test are summarized as follow: The blind *S*s made on an average 2.2 errors, the sighted *S*s 0.9 errors; the early blind 2.9; and the late blind 1.4. Using the *t*-test the difference between the blind and the sighted, the early and the late blind, and the early blind and the sighted are all significant at the 1-% level. The smallness of the difference between the late blind and the sighted arises from the fact that the test was rather too easy for this whole sample, and even among the early blind there were some who made maximal scores. In Worchel's investigation all the differences were significant, and by combining the two sets of data we have results which are very solid indeed. It should be pointed out in this connection that Worchel's division between early and late blind was made at 6 yr. of age. Four of our present group of late blind would thus have been included in his early blind category. The figures have been worked over using his limits and the differences remain.

No significant figures emerge when performance on the test is related to age or intelligence, or when performance by the late blind group is related to time elapsed since blinding.

(2) Orientation

(a) Method. The problems of orientation were presented by means of a peg-board one foot square with 10 rows and columns spaced 1 in. apart (Fig. 1). In these holes could be placed pegs fitting easily and projecting ¼ in. from the surface of the board. The board was placed squarely on the *S*'s lap as he sat leaning slightly backwards and with head supported. His left hand was so placed along the left hand edge of the board that he could establish its position. *E* then took the *S*'s right hand and moved it at a constant speed from one peg to another touching each peg with

the forefinger. The positions of the pegs are shown in Table I. They included two pegs placed vertically, horizontally, and diagonally; three pegs making right angles, obtuse angles, and acute angles; and four pegs making a square and a diamond. Ten problems were given in all. At the end of each sequence *S* was told to grasp the last peg between finger and thumb while the board was rotated through 180° and the other peg or pegs removed. He was then given these and told to replace them in

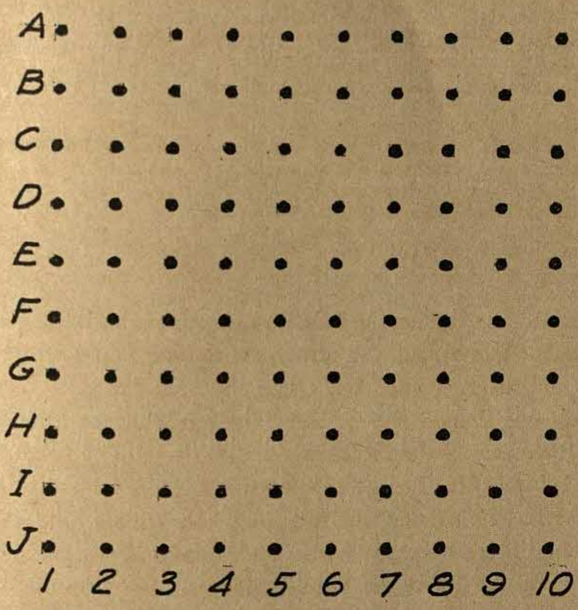


FIG. 1. PEG BOARD USED IN ORIENTATION-PROBLEMS

Rows and columns spaced at intervals of 1 in.

the holes they had occupied when he touched them first. Errors were measured in inches from the correct hole, and a low score thus represents a good performance.

(b) Results. The results of the orientation-test are as follows: Blind ($N = 37$) mean score, 33.1; Sighted ($N = 37$), 41.2; Early blind ($N = 19$), 39.6; Late blind ($N = 18$), 26.2. Using the *t*-test for significance the difference between the late blind group and the sighted group is significant at the 1-% level, as is the difference between the late and early blind. The difference between the blind and the sighted is significant at the 5-% level. The difference between the early blind and the sighted is not significant.

We thus have a task which the blind on the whole do better than the sighted, but one in which the late blind do conspicuously better than either

the early blind or the sighted and in fact contribute most of the difference between the blind and the sighted groups.

An attempt to correlate performance with age or intelligence gave no significant figures, nor was there any relationship in the late blind group between performance and time elapsed since blinding. Though they might not have been predicted from earlier work on orientation, these results are reasonable enough. The task is one which involves at the same time the use of tactile-kinesthetic cues and an ability to deal with a spatial pattern which moves, as it were, all of a piece. The early blind have been long

TABLE I

Position of the Pegs in the Orientation-Problems

Problems

2 peg	3 peg	4 peg
(1) B5, H5	(4) B8, H8, H2	(9) C7, G7, G3, C3
(2) E9, E3	(5) H8, B8, H2	(10) F7, J5, F3, C5
(3) C8, I3	(6) G9, C6, G3	
	(7) C7, H5, G2	
	(8) F6, C7, J2	

dependent upon tactile-kinesthetic cues and are presumably skilled in their use. Nevertheless some teachers of the blind have claimed that early blind children never equal late blind children in many structural-manipulative tasks as well as having the sort of trouble one might expect with subjects like geometry. There seems in fact to be a generalized defect of space perception associated with early blindness, and this has shown up here as well as in some of the investigations to which reference has already been made.

Turning to the sighted *S*s, who after all did rather less well than the early blind, we have clearly to look for quite a different kind of explanation. The obvious one is tactile-kinesthetic inexperience leading to a failure to coördinate the tactile-kinesthetic cues with such visual imagery as may have been present. There is some additional evidence to support this. The behavior of the sighted *S*s was generally more groping and uncertain and on some of the problems they were more likely than the blind to rotate their reconstruction of the pattern through 90° or 270°. The frequency of this error among the blind as compared with the sighted on Problems 4 and 5 was 17 to 32. This difference is significant at the 1-% level.

In contrast with both early blind and sighted the late blind group seem to have on the one hand the kind of space perception that enables them to keep track of a 180° rotation, and on the other the tactile-kinesthetic skill to deal with the distances and directions on the board in front of them.

It may not be obvious from the figures how high a level of skill was shown in a good late blind performance. A total error of 7.2 in. for all 10 problems was the best. No sighted *Ss*, even sophisticated *Ss* such as research students and staff have been able to get below 20 in. In the experiment itself only one early blind *S* was better than 20 in. as against five late blind *Ss*.

It may seem that the evidence presented in this experiment indicates that late learning can be more important than early learning. Vision at 5 or 6 yr. of age seems crucial in a way that vision at 2 or 3 yr. is not. Though this interpretation cannot be ruled out it is almost certainly an oversimplification. What is more probable is that some of the visual and other skills involved in space perception are so complex that they require a long period of training and a mastery of the simpler skills rather than a nervous system at a certain level of maturity. It will be difficult ever to discriminate between these two possibilities if indeed they are genuine alternatives.

(3) Classification

(a) Method. The classification-test was relatively brief and simple. As all the tests were given at the same session with the *Ss* somewhat restricted in their move-

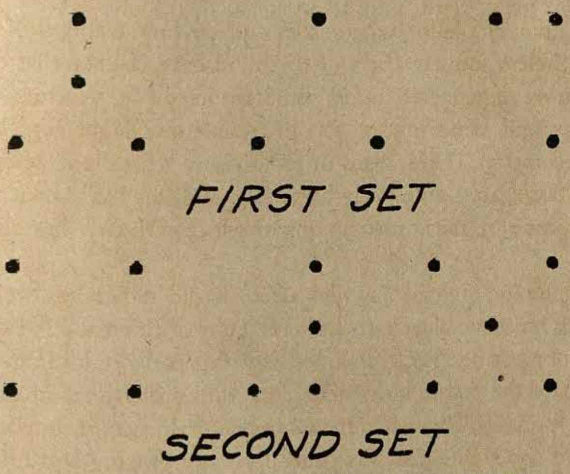


FIG. 2. FIGURES USED IN THE CLASSIFICATION-TEST

ments because of the *EEG* equipment it was desirable to keep testing time down to between 30–40 min. On a peg board wider than that used in the orientation test, but with the holes again 1 in. apart, two sets of three figures each were outlined using three, four, or five pegs (Fig. 2). *S* was instructed as follows.

On the board in front of you there are three figures outlined by the pegs. [Taking *S*'s right hand, *E* continued:] Here is one, and another, and another. They are all different, but two of them are alike in some way and the other is 'odd man out.' I want you to tell me which two you think are alike and why.

Set 1 and Set 2 were each presented first an equal number of times. It can be seen that several criteria are possible such as shape, number of pegs, no peg in middle, symmetry, and so on. In scoring only the shape and counting criteria were used leaving a miscellaneous group which fortunately was almost the same size for both blind and sighted.

(b) *Results.* Results are as follows: blind *S*s, shape 40, counting 17, miscellaneous 17; blindfolded sighted *S*s, shape 24, counting 31, miscellaneous 19. If one disregards the miscellaneous column these figures are significant at the 1-% level. No significant differences were obtained between the two blind groups.

As a supplementary experiment the two sets were presented visually to 160 *S*s. Of these, 32 *S*s used shape as the basis for their classification and only 10 *S*s used counting. In other words if we compare tactile-kinesthetic classification of blind *S*s with that of sighted *S*s we find that the blind approximate more closely to a classification based upon seeing. This is what might have been expected. The blind person has to get to know his world by way of touch and movement, and though early blindness may, as we have seen, impose some limitations upon him, his ability to gain information from his hands is likely to be higher than in the case of sighted persons. In particular he should be better at translating pressure and movement into the linguistic categories used in everyday life. That is what he has had to do since he became blind and that is what he was doing in this test.

Hunter's Experiment

Perception of straightness. Another aspect of the superiority of the blind in making use of tactile-kinesthetic information showed up in Hunter's experiment with some of the same *S*s.[12] He presented them with a flexible steel ruler clamped at the ends and with curvature controlled by a micrometer screw at the mid-point. *S* ran his fingers along the top edge and had to judge whether the ruler was straight, bent away, or bent towards. While apparent straightness corresponded to a curve away for both blind and sighted *S*s, the mean curve of the blind *S*s was nearer objective straightness. The blind also tended to vary less in their judgments. Both differences were significant at the 1-% level. This superiority in the blind was already present by the age of 13 yr. and the later age-groups were no more objective though their variability was less.

Data from this and from the classification experiments clearly strengthen the suggestion that the superiority of the blind over the sighted in the orientation-test is a matter of tactile-kinesthetic skills, and that the marked superiority of the late over the early blind is due to the presence of an additional visual component. This last would appear to be a complex cumulative ability of some sort which requires 5 or 6 yr. for its development.

---

[12] I. M. L. Hunter, Tactile-kinesthetic perception of straightness in blind and sighted humans, *Quar. J. Exper. Psychol.*, 1954, 149-154.

Conclusions

Nothing has been presented here which supports Hebb's suggestion about the distinction between early and late learning so directly as does the work of Senden and Riesen. Even though the learning upon which later performance was shown to depend happened to be early learning we have not shown that this must be so. At the same time indirect support is not lacking. It has been established that certain perceptual abilities having to do with objects in space seem to require a long apprenticeship either in the visual or in the tactile-kinesthetic modalities, and that once this apprenticeship has been served different amounts of later practice have no appreciable effect. We have in fact something rather like the kind of abilities identified by factorial studies of test-performance.

A more general and perhaps obvious point seems relevant to the experimental study of learning and perception. The attempt to establish theories of wide generality upon the performance of adult *S*s, who are called naïve merely because they have never taken part in a similar experiment, has risks which are plain enough even though they may be ignored. To say, when such a *S* learns something, that this is the way he always learns is like deriving the laws of building construction from watching men put a few tiles on the roof. The same holds good in the field of visual perception. Much of the work here is simply a study of how expert perceivers adjust to systematic manipulation of some basic visual cues, and there is no direct way from this to a knowledge of how we learn to perceive.

The experimental study of learning and perception can justify itself on its own terms; but if, as seems reasonable, the naïve *S* is a sort of hypothetical construct then we need some reliable bridge between him and the flesh and blood *S*s upon whom we experiment. In other words it would appear that genetic psychology is not a separate field but an integral part of any experimental study of high level functions. This, of course, is implied by Hebb's distinction between early and late learning.

Summary

Seventy four *S*s divided into matched groups of early blind, late blind, and sighted were given three spatial tests.

(1) A figure-recognition test showed the sighted to be slightly superior to the late blind and the late blind to be much superior to the early blind.

(2) A test which involved the ability to handle a 180° rotation within the perceptual field showed the late blind to be considerably superior with the early blind and sighted very close to one another.

(3) A short classification-test in the tactile-kinesthetic modality showed

that the blind may approximate more closely to the visual performance of the sighted than do the sighted themselves.

(4) A test of the tactile-kinesthetic perception of straightness showed that the blind are superior to the sighted both in objectivity and consistency.

In none of these tests did it appear that the differences found were due to differences in the time between the onset of blindness and the investigation provided that the onset occurred after the age of 4 yr. and not less than 2 yr. before the tests were taken.

Age and intelligence were not found to have significant effects within the range studied.

The findings seem to point to the existence of certain basic skills which are built up over a period of years, and underlie performance in ways not unlike those suggested for abilities by workers in the field of mental testing. In the cases studied these skills seem to have been built up early and later learning has little effect.

THE EFFECT OF ANGLE OF SLANT UPON THE TRIGONOMETRIC RELATIONSHIP OF PRECISION AND ANGLE OF LINEAR PURSUIT-MOVEMENTS

By RICHARD F. THOMPSON and W. J. BROGDEN,
University of Wisconsin

Slant is an inclination of a surface from the horizontal position in the plane parallel to the frontal surface of the observer. This paper presents the results of an experiment on the effect of angle of slant upon the trignometric relationship of precision and angle of linear pursuit-movements. The equation for the trigonometric relationship is given by Corrigan and Brogden as follows:[1] $y = a + b \cos 2x + c \sin 2x$, in which y = precision of right-arm movements in terms of group-mean frequency of stylus-contact, x = angle from the body at which the movement is made,[2] a = the constant that determines the baseline of the curve (mean frequency of stylus-contact for all angles), and b and c are the constants by means of which the amplitude (d) and the phase angle ($2e$) of the curve are determined [$d = (b^2 + c^2)^{1/2}$; $\cos 2e = c/d$].

Briggs, Thompson, and Brogden studied the effect of angle of tilt upon the trigonometric relationship of precision and angle of linear pursuit-movements.[3] Tilt was defined by them as inclination of the tracking surface from the horizontal position in the plane perpendicular to the frontal plane of *S*. The trigonometric equation proved to be an excellent fit of the relation of precision and angle of linear pursuit-movements for the data obtained at each of the angles of tilt tested by the experiment. The fitted equations, however, differed as a function of angle of tilt. Since angle of slant is a variable similar to angle of tilt, determination of its effect upon the trigonometric relationship of precision and angle of linear pursuit-movements is the purpose of the present investigation.

---

* Accepted for publication October 13, 1954. Supported in part by the Research Committee of the Graduate School from funds granted by the Wisconsin Alumni Research Foundation.

[1] R. E. Corrigan and W. J. Brogden, The effect of angle upon precision of linear pursuit movement, this JOURNAL, 61, 1948, 502-510; Corrigan and Brogden, The trigonometric relationship of precision and angle of linear pursuit-movements, *ibid.*, 62, 1949, 90-98.

[2] Angle is designated by Cartesian coördinates. 0° represents the position in which the track is normal to the frontal plane of the body of *S*, and the arm-movement is started with the stylus close to, and continued away from the frontal plane of the body.

[3] G. E. Briggs, R. F. Thompson, and W. J. Brogden, The effect of angle of tilt upon the trigonometric relationship of precision and angle of linear pursuit-movements, this JOURNAL, 67, 1954, 475-483.

Method and Procedure

Apparatus. The apparatus is the same as that used in previously reported experiments.[4] It consists of a track formed by two brass plates resting on a piece of glass that *S* traverses with a metal-tipped stylus. Velocity of stylus-movement is controlled by instructions given *S* to match the rate of his movement to that of a small cylindrical target that travels beneath the glass plate of the track at a constant velocity of 3.0 cm. per sec. Control devices automatically start the target, stop it at the end of the track, and return it to the starting position where its direction is reversed for the start of a new trial. The platform on which the track is mounted may be rotated about its center. The angle 0° is represented by the track normal to, and the long axis of the target parallel to, the frontal plane of *S*. A disk attached to the bearing of the central vertical axis is notched every 15° to permit selection of the desired angle. The platform may be rotated also in its vertical plane parallel to the frontal plane of *S* to provide for adjustment of angle of slant. Angles produced by inclination of the platform toward the left of *S* are negative in sign and those toward the right are positive. A milled index permits selection of the angle of slant desired. The stylus and each side of the track constitute an open switch in parallel with the counting switch of a Potter Electric Counter (Model 67). Each contact of the stylus with a track side is registered cumulatively on the counter. When the target is reversed at the end of the track, the input to the counter is closed so that no further recording may occur. During the time the target is being reversed, *E* records the number of errors for the trial just completed. When the target reaches the head of the track, the action reversing its direction also clears the counter which then registers zero at the start of the next trial.

Subjects. Eighty students, men who were right-handed, served as volunteer *Ss* in the experiment. None had previous experience in this experimental situation.

Procedure. *S* was seated in a tank-driver's seat and *E* adjusted a harness about *S*'s trunk to maintain the right shoulder in a fixed position. The height of the seat and its closeness to the track-platform were so adjusted to his body stature that *S* was both comfortable and able to reach the end of the track when his right arm was fully extended. The standardized instructions, reported verbatim in a previous paper,[5] were modified slightly as required by the introduction of angle of slant as an experimental variable.

A randomized 8 × 8 greco-latin square was employed to determine the combination of angle of slant and angle from the body, the ordinal position of the combinations, and the sequences of the combinations of the experimental design. Each angle of slant and each angle from the body occurs once and only once in each row and column of the square. The combination of angles differs in each row of the square. The angles of slant are −30°, −20°, −10°, 0°, 10°, 20°, 30°, and 40°. Angle from the body is represented by values of 0°, 30°, 45°, 60°, 90°, 120°, 135°, and 150°. The 80 *Ss* were assigned at random to the rows of the square, with 10 *Ss* completing each of the 8 sequences. Thus, there were 10 replications of the square.[6] *S* completed three practice-trials at the initial combination of angles of the sequence to

[4] Corrigan and Brogden, *opp. citt.*, 502 f and 90 f.

[5] Corrigan and Brogden, *op. cit.*, 502 f.

[6] Davis Chambliss served as *E* in work with 16 *Ss* (two replications of the square), for which the authors express their appreciation.

which he was assigned. Thirty seconds after practice was completed, he received 10 trials at each of the 8 combinations of angles for his sequence (row of the square) with a 30-sec. rest between each block of 10 trials.

RESULTS

Transformation of the error-scores is the first step in the treatment of the data. The natural logarithm of the *raw score* + *5* was obtained for every trial for every *S*. The means for the 10 trial-blocks in each cell of the rows of the greco-latin square were computed for every *S*, and constituted the basic data for the analysis of variance. The procedure for analysis followed that given by Edwards for replicated latin squares.[7] A summary of this analysis is given in Table I. Separate *F*-tests, using the 5-% level of

TABLE I

SUMMARY OF RESULTS OF ANALYSIS OF VARIANCE

Source	df.	Sum of squares	Mean square	F	F at 5-% level
Total between Ss	79	2417.36			
Sequences	7	142.60	20.37	0.64	2.91
Residual between Ss	72	2274.76	31.59		
Total within Ss	560	4157.34			
Angle	7	1259.30	179.90	14.31	2.24
Slant	7	494.78	70.68	5.62	2.24
Ordinal position	7	301.89	43.13	3.43	2.24
Square uniqueness	42	527.84	12.57	3.97	1.45
Residual within Ss	497	1573.53	3.17		
Grand total	639	6574.70			

confidence, were made for the uncorrelated data, between-*S*s variation, and for the correlated data, within-*S*s variation. The single test of sequences for the uncorrelated data is not significant. For the correlated data, the square-uniqueness mean square is significant when tested against the residual within-*S*s mean square. The square-uniqueness mean square must, therefore, be used as the denominator in *F*-tests for the angle, slant, and ordinal-position mean squares. All of these *F*-tests are significant. Further analysis of the means for angle, slant, and ordinal position are therefore appropriate.

Examination of the means for ordinal position shows a progressive decrease in the magnitude of the error-score with increase in ordinal position. Such a practice effect has been commonly found in related studies of linear pursuit-movements.[8] Separate analyses of variance were completed on

[7] A. L. Edwards, *Experimental Design in Psychological Research*, 1950, 446.

[8] Corrigan and Brogden, *op. cit.*, 90 f; G. E. Briggs and W. J. Brogden, Bilateral aspects of the trigonometric relationship of precision and angle of linear pursuit-movements, this JOURNAL, 66, 1953, 472-478.

the data for each one of the eight angles of slant. Significant F-ratios for slant angles of 0°, 10°, 20°, 30°, and 40° were obtained. Thus, there is evidence of an interaction between angle of slant and angle from the body. Variation of angle from the body produces significant differences in precision of linear pursuit-movements at slant angles of 0°, 10°, 20°, 30°, and 40°, but not for the negative angles of −10°, −20°, and −30°. Duncan's range test was applied to the data for the angles of slant for which significant F-ratios were obtained.[9] The results of these tests indicate that for each set of data there are at least two groups of scores (*e.g.* most difficult angles and least difficult angles) which differ significantly from each other at the 5-% level of confidence. These results justify fitting the trigonometric equation to the data for slantangles of 0°, 10°, 20°, 30°, and 40°. The equation was fitted also to the data for the negative angles of slant in the attempt to obtain a complete analysis of the interaction of angle of slant and angle from the body, in spite of the failure to obtain significant differences.

The least-square regression equations of the trigonometric function, $y = a + b \cos 2x + c \sin 2x$, for the eight angles of slant are plotted in Fig. 1 together with the empirical data. The Chi-square test of goodness of fit yielded Chi-squares, for all curves, that are not significant at the 99-% level of probability. The trigonometric equation, therefore, provides adequately for the mathematical expression of the relationship between precision and angle of linear pursuit-movements at each one of the eight angles of slant. These results provide justification for including the equations for the negative angles of slant in further examination of the effect of angle of slant upon the relationship of precision and angle of linear pursuit-movements.

Fig. 1 shows a progressive change in the curves of precision and angle of linear pursuit-movements as the angle of slant is increased by 10° steps from −30° to 40°. These changes are shown more precisely by Fig. 2. Curve A is a linear plot of the curves of Fig. 1 against angle of slant. Curves B, C, and D are plots of the phase angle, baseline, and amplitude respectively, against angle of slant. These latter curves express concisely the complex effect of angle of slant upon the relationship of precision and angle of linear pursuit-movements shown grossly by curve A.

Phase angle decreases in magnitude as angle of slant increases. It is positive for the negative angles of slant and for 0° slant, and is 0° or negative for slant angles of 10° to 40°. The form of this curve may be that of

[9] R. B. Duncan, A significance test for differences between ranked treatments in an analysis of variance, *Virginia J. Sci.*, 2, 1951, 171-189.

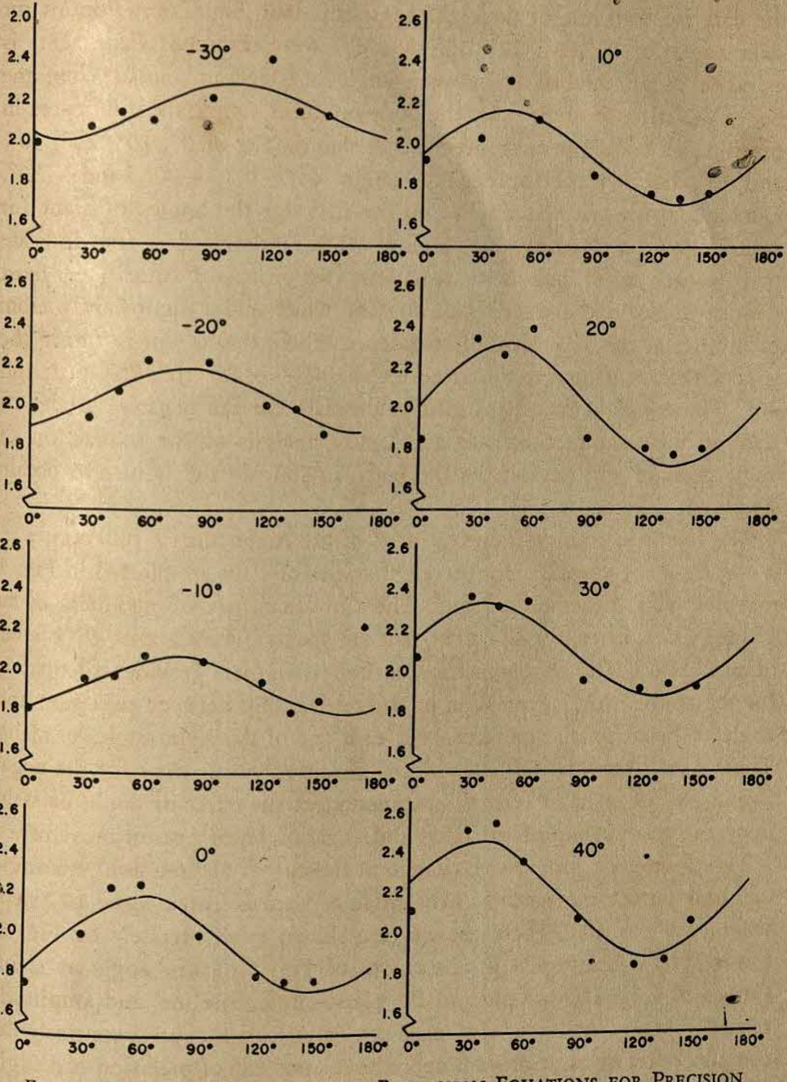


FIG. 1. CURVES OF TRIGONOMETRIC REGRESSION EQUATIONS FOR PRECISION AND ANGLE OF LINEAR PURSUIT-MOVEMENTS FOR ANGLE OF SLANT

For each curve, the ordinate is the logarithm (base e) of the *raw error score* + 5 and the abscissa represents the angle from the body at which performance was measured. Every point is the group-mean of the *Ss'* mean precision of performance for the block of 10 trials at the angle indicated. The curves are plots of the equation, $y = a + b \cos 2x + c \sin 2x$, fitted to the empirical data by the method of least-squares.

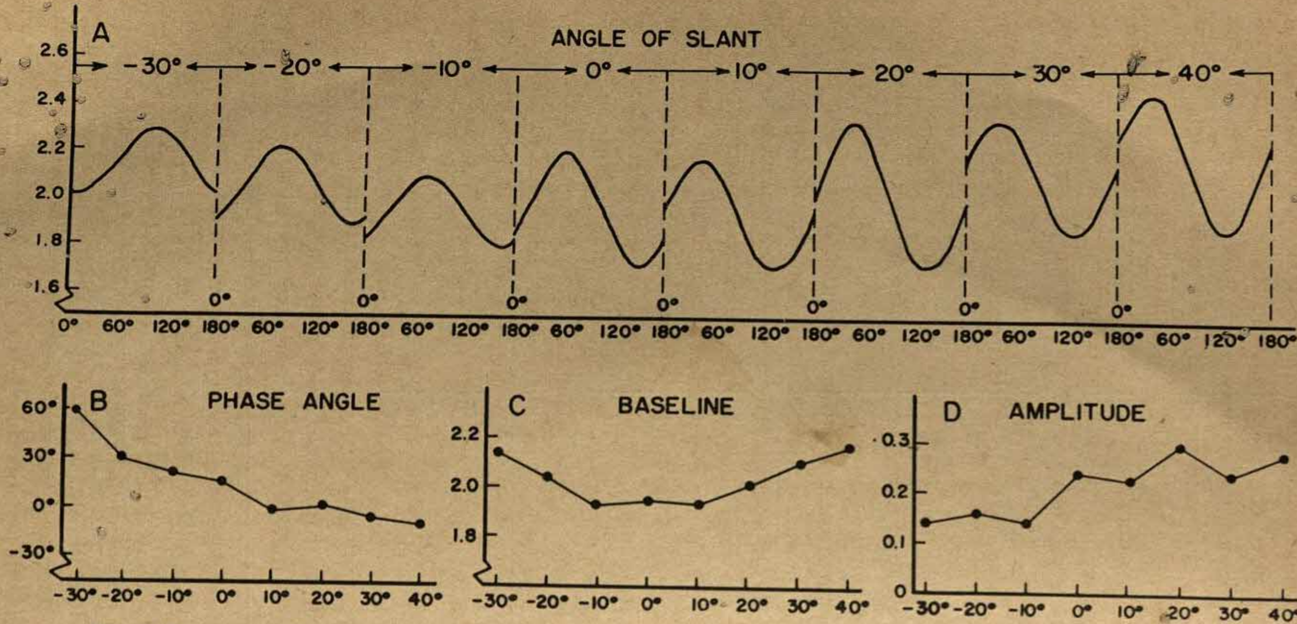


FIG. 2. REGRESSION EQUATIONS, PHASE ANGLE, BASELINE, AND AMPLITUDE AS A FUNCTION OF ANGLE OF SLANT

Curve *A* is a linear plot of the regression equations of Fig. 1 against angle of slant. Curves *B*, *C*, and *D* are plots of characteristics of the regression equations against angle of slant. The ordinate is the transformed error-score except for phase angle, where it is degrees. The abscissa is angle of slant, except for *A* where both angle from the body at which the movement is made and angle of slant are represented.

an exponential decay-function or of a linear function. Further curve fitting of data derived from fitted regression equations is not, however, justified. The significance of the decrease in phase angle with increase in angle of slant is the change in relative difficulty of linear pursuit-movements as a function of angle from the body. The most difficult angle at −30° slant is on the order of 100°. As slant-angle is increased, the most difficult angle decreases in magnitude until at 40° slant, it is on the order of 40°. There is, of course, a concomitant change in the magnitude of the least difficult angle, which is on the order of 10° at −30° slant, 160° at −20° slant, with general decrements for each angle of slant up to 40°, where it is on the order of 120°.

Curve *C*, showing the relation between magnitude of the baseline and angle of slant, first shows a decrement as angle of slant increases, then reaches a minimal magnitude for angles of slant of −10°, 0°, and 10°, and thereafter increases progressively with increase in angle of slant up to 40°. The baseline is the mean precision of linear pursuit-movements over all angles from the body. If its magnitude is small, then general precision of movement is high; if large, the general precision of movement is low. The baseline, however, is not independent of amplitude, since amplitude represents the magnitude of the difference in precision of performance at the most and least difficult angles. Curve *D* represents this latter relation, and although it has several inversions, the tendency appears to be one of increase in magnitude with increase in angle of slant. Although baseline is at minimum for slant-angles of −10°, 0° and 10°, amplitude for the slant-angle of −10° is markedly less than that for 0° and 10°. On the other hand, maximal precision for some angles from the body occurs at slant-angles of 0°, 10°, and 20°.

Discussion

Since phase angle, baseline, and amplitude are characteristics of the trigonometric equation that are not independent of each other, there is no simple expression of the effect of angle of slant upon the relationship of precision and angle of linear pursuit-movements. Each of these characteristics undergoes a progressive but different change as angle of slant is increased. These functions represent changes in the fitted regression equations that are readily discernible when the curves of the equations are presented on a linear plot of angle of slant.

The effect of angle of slant upon the relationship of precision and angle of linear pursuit-movements is both similar to and different from the effect of angle of tilt upon this relationship.[10] The planes of rotation for slant and

[10] Briggs, Thompson, and Brogden, *op. cit.*, 475 f.

tilt are at right angles to each other. Under the conditions of the two studies, with the right shoulder of *S* held in a fixed position, the sectors of the two planes of rotation available for experimental manipulation are different. For slant, the sector is between angles of −30° and 40°, whereas for tilt the sector is between 0° and 70°. The limits of these sectors represent roughly the extremes of tilt and slant at which *S*s can perform linear pursuit-movements. For both tilt and slant the range is 70°, but the orientation with respect to the horizontal position of the platform, 0°, is different for the two.

Neither the effect of slant nor the effect of tilt upon the trigonometric relationship of precision and angle of linear pursuit-movements may be given simple expression. There is, however, a progressive alteration of the form of the fitted regression equations in both cases, when they are plotted as a linear function of angle of slant or of tilt. In the case of tilt, increasing the magnitude of the angle produces functions of phase angle, baseline, and amplitude that are much less regular than those for angle of slant. The curve of phase angle and tilt appears to present two levels of magnitude of phase angle for two sets of adjacent angles of tilt. One level is given by the mean of 17.8° for angles of 0°, 10°, 20°, and 30°; the other by the mean −20.0° for angles of tilt for 40°, 50°, 60°, and 70°. For baseline and amplitude, there are progressive changes in magnitude as angle of tilt is increased. Although the two curves are different from each other, both show decreases in magnitude with increase in angle of tilt between 0° and 20°, remain at a minimum between 20° and 50°, and then increase progressively at 60° and 70°. Thus, only the curves for baseline show similarity for angle of tilt and angle of slant.

The results of the present paper support the view expressed by Briggs, Thompson, and Brogden that angle from the body is a primary determiner of precision of linear pursuit-movements and that the integrity of the trigonometric relationship of precision and angle of linear pursuit-movements is maintained when other variables are manipulated.[11] Angle of slant is another variable subsidiary to angle from the body. The trigonometric relationship is altered, but not destroyed by any of the angles of slant used in the experiment.

Summary

An 8 × 8 greco-latin square was employed to study the effect of angle of slant upon the trigonometric relationship of precision and angle of linear pursuit-movements. The angles of slant were −30°, −20°, −10°, 0°, 10°,

[11] *Loc. cit.*

20°, 30°, and 40°. Angles from the body were 0°, 30°, 45°, 60°, 90°, 120°, 135°, and 150°. Analysis of variance indicated that angle of slant, angle from the body, and ordinal position were significant sources of variation. The trigonometric equation, $y = a + b \cos 2x + c \sin 2x$, was fitted by the method of least-squares to the data for each angle of slant, and all fits were found to be good by the Chi-square test. The effect of angle of slant upon the relation of precision and angle of linear pursuit-movements is expressed by curves of phase angle, baseline, and amplitude, each plotted against angle of slant. Phase angle decreases as angle of slant increases. Baseline decreases to a minimum and then increases as angle of slant increases. Amplitude increases as angle of slant is increased. The effect of slant is complex, but it is subsidiary to angle from the body as a determiner of precision of linear pursuit-movements.

RESPONSE-SEQUENCES IN A PAIR OF TWO-CHOICE PROBABILITY SITUATIONS

By JACQUELINE JARRETT GOODNOW, Walter Reed Army Institute of Research

In situations where a person has to predict, over a number of occasions, which of two alternative events will occur, several experimenters have observed tendencies to alternate predictions or to anticipate Event B after a run of Event A. For the sake of convenience, we shall refer to such tendencies as recency-effects.[1] The analysis presented here asks whether such effects may be regarded as a function of the particular stimulus-material and the way the problem is presented to *S*.

More specifically, the following hypotheses are tested: (1) When the task is presented to *S* as one involving 'chance,' and when the only difference between one event and another is their position in a series, recency-effects based on *S*'s notions about runs of events will occur. This is the kind of task in which recency-effects have been noted, *e.g.* *S* predicts whether or not a light will flash on or whether *E* will say 'plus' or 'check;'[2] (2) When the task is presented to *S* as one not involving 'chance,' and when a number of aspects of the events vary from trial to trial, such recency-effects will not occur, even though the sequence of events in the two tasks is the same.

The major interest of recency-effects lies in the questions they raise about the probability of a response or prediction occurring on a particular trial as compared with the probability of a response occurring over a series of trials. If, for example, a series of trials is so arranged that Event A occurs on a random 50% of the trials and Event B occurs on the other 50% of the trials, it has been found that *S*s learn to anticipate Event A on 50% of their choices over the series and Event B on the remaining 50% of their choices.[3] The question arises of whether the same probabilities

* Accepted for publication, November 22, 1954. The experiments reported were conducted in 1951 at Harvard University and supported by the Laboratory of Social Relations. The author wishes to express appreciation of advice given by Dr. Frederick Mosteller and Dr. Robert Bush.

[1] A recent review of such effects is given by V. L. Senders and Ann Sowards, Analysis of response sequences in the setting of a psycho-physical experiment, this JOURNAL, 65, 1952, 358-374.

[2] H. W. Hake and Ray Hyman, Perception of the statistical structure of a random series of binary symbols, *J. Exper. Psychol.*, 45, 1953, 64-74. M. E. Jarvik, Probability learning and a negative recency effect in the serial anticipation of alternative symbols, *J. Exper. Psychol.*, 41, 1951, 291-297.

[3] D. A. Grant, H. W. Hake, and J. P. Hornseth, Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement, *J. Exper. Psychol.*, 42, 1941, 1-5; Hake and Hyman, 68; Jarvik, 292; W. O. Jenkins and J. C. Stanley, Partial reinforcement: A review and critique, *Psychol. Bull.*, 41, 1950, 193-234.

hold for responses occurring on particular trials. In a simple form, the question may be asked whether, following the occurrence of Event A on three previous trials, the probability of Event B being anticipated on the next trial is 0.5 or significantly different from 0.5. Where negative recency-effects operate—*e.g.*, where there is a tendency to anticipate Event B after a run of Event A—the anticipation of a further occurrence of Event A will clearly depart from a probability of 0.5. Such departures and the form they take are of interest in any attempt to predict responses on particular trials. They have also figured largely in discussions of the hypotheses *S* uses in making predictions and of the extent to which recency-effects explain the finding that over a series of trials *Ss'* choices match the event-probabilities.

DATA FOR ANALYSIS

Data were provided by an experiment comparing performance in a matched pair of two-choice tasks, one a gambling task and the other a problem-solving task.[4] These tasks had been designed to parallel one another in all but two important aspects—the gambling task involved chance and the stimulus-objects varied only in temporal position in the series; the problem-solving task involved searching for a logical principle and the stimulus-material took a number of different but recurring forms.

Gambling task. In the gambling task, *S* played a kind of one-armed bandit. The immediate stimulus was the flashing on of a signal light to indicate that the device was ready to take a bet. The two alternative responses were pressing a right key and pressing a left key, the key pressed being the one on which *S* placed his 1¢ bet. The amount to be won or lost on any one trial was always the same—1¢.

Problem-solving task. In the problem-solving task, *S* worked on an insoluble problem involving a variety of geometrical designs. On each trial, *S* was presented with three cards. The first of these contained a key geometrical design. The other two cards contained variations on the key design. The variations presented were always of two types, formed by adding or subtracting a line to the key design. On each trial, *S* coupled the key design with one of its variations and sought to find the principle by which the coupling should be made. The immediate stimuli then were the three cards presented. Variation in stimulus-material arose from the use of 10 different key designs with 4 variations on each key design. The two alternative responses were (1) choice of the subtracting variation and (2) choice of the adding variation to couple with the key card.

The two tasks, gambling and problem-solving, were matched in the following respects:

Event-probabilities. In both tasks the event-probabilities were 50:50, 70:30, and 90:10. Event-probabilities refer to the frequency with which each of the two alternatives was scheduled to pay off. Case 70:30 with the gambling task, for instance, meant that one key was scheduled to pay off on a random 70% of the trials and the other key on the remaining 30%. *S* learned which key was scheduled to pay off on any trial only by pressing one key and winning or by pressing the

[4] J. J. Goodnow, Determinants of choice-distribution in two-choice situations, this JOURNAL, 68, 1955, 106-116.

other key and losing. He was told, however, that one and only one alternative could be the correct choice on any one trial.

Payoff schedule. The same schedules of event occurrence were used for both tasks. For all *S*s in the 70:30 and 90:10 groups, the same event was given the higher probability of occurrence—the subtracting variation in the problem-solving groups and the left key in the gambling groups were more often correct. Randomization of event-occurrence was within blocks of 10 trials.

Arrangement of trials. All *S*s were allowed 120 trials, divided into 12 blocks of 10 trials each. Inter-trial and inter-block intervals were respectively 3 min. and 10 min. For the problem-solving task, each of the 10 key cards was presented once in each block of trials. *S*'s choices in both tasks were recorded by *E*.

Subjects. *S*s were Harvard undergraduates, obtained from the Student Employment Bureau and paid 75¢ an hour. From 10 to 14 *S*s were run in each of the 6 groups, each *S* in a group being given one of the three probability ratios and one of the two tasks. Both tasks then provided a choice-by-choice record over 120 trials for each *S*.

Not all the records were used for analysis, as *S*s who distributed choices in 100:0 or close to 100:0 fashion were excluded. These *S*s, choosing the one alternative so consistently, provide little or no data on different kinds of response-sequence, and their exclusion left us, in both tasks, with *S*s whose choice-distributions over the series of trials were not significantly different from the event-probabilities. The groups most affected by this exclusion were the 90:10 groups. Records could be used of 5 *S*s in the 90:10 problem-solving group and 3 *S*s in the 90:10 gambling group. In the 70:30 gambling group, records for 8 *S*s could be used. Accordingly, for reasonable comparability of data, in each of the remaining three groups (both 50:50 groups and the 70:30 problem-solving group) the first 8 *S*s were taken whose choice-distributions were not significantly different from the event-probabilities.

Method of Analysis

It would be possible to determine for each trial or for certain single trials the proportion of *S*s choosing each of the two alternatives. This would require, however, a large number of *S*s. The usual procedure is to pool a number of trials which can be regarded as equivalent and observe the proportions of *S*s choosing each of the alternatives on these pooled trials.

The procedure used, described in terms of the gambling task, is as follows. We distinguish between four kinds of preceding trials in terms of the choice made (left or right key) and its outcome of win or loss. We then determine for each of these groups of trials the number of times that each kind of trial is followed by a choice of left, and the number of times it is followed by a choice of right. This count is made over trials 21-120. The first two blocks are excluded because the distributions of responses, reflecting *S*s' initial preferences, are highly variable at this stage.

The 100 trials remaining yield 90 response pairs; the sequence between the last response in a block of 10 trials and the first response on the next block is not counted because of the longer inter-block time interval.

The incidence of left and right choices after a particular kind of trial is compared with the overall incidence of left and right choices on the second member of pairs

of responses. We may find, for instance, that, in the gambling 50:50 group, choices on trials immediately after an unsuccessful choice of left consist of 68% left choices against 32% right choices. In contrast, total second choices are made up of 47% left choices against 53% right choices. Clearly, after an unsuccessful choice of left the occurrence of a further choice of left is higher than we would expect from the general frequency of left and right choices.

RESULTS

Table I shows, for the two tasks, the incidence of left and right choices or their equivalents following the four kinds of preceding trial.

As a summary statement before discussing the results in detail, the

TABLE I

THE INCIDENCE OF CHOICES OF LEFT AND RIGHT ALTERNATIVES OR THEIR EQUIVALENTS, AND THE PROPORTION OF LEFT CHOICES, FOLLOWING EACH OF FOUR KINDS OF PRECEDING TRIAL

Groups differ in event probabilities (50:50, 70:30, and 90:10), and in task (Task A is Problem-solving, Task B Gambling.)

Type of Preceding Trial and Subsequent Choice

Group	Chose L (won) next choice			Chose L (lost) next choice			Chose R (won) next choice			Chose R (lost) next choice			Total choices		
	L	%L	R	L	%L	R	L	%L	R	L	%L	R	L	%L	R
50:50 Task A	80	48%	85	86	46%	100	85	47%	95	88	46%	101	339	47%	381
50:50 Task B	89	48%	95	109	68%	52†	91	45%	112	49	28%	123†	338	47%	382
70:30 Task A	177	62%	108	71	58%	51	44	49%	46	120	54%	103	412	57%	308
70:30 Task B	219	69%	101	90	69%	41	44	52%	41	95	52%	89*	448	62%	272
90:10 Task A	304	88%	45	35	88%	5	5	100%	0	45	80%	11	389	86%	61
90:10 Task B	172	86%	27	15	75%	5	7	100%	0	24	54%	20†	218	87%	52

* Significant at 1-% level.
† Significant at 0.1-% level. Significance levels were established with binomial probability paper.[5]

following points can be made: (a) In the groups given the *problem-solving* task, the ratios of subtracting to adding choices following each of the four kinds of preceding trials are not significantly different from the ratio of total subtracting to total adding choices. (b) In the groups given the *gambling* task, the ratio of left to right choices following certain kinds of preceding trial does significantly depart from the ratio of total left to total right choices, in a way which indicates that negative recency-effects are operating.

In the 50:50 gambling group, *S*s tend to repeat their choice of a key after choosing this key on the previous trial and losing. This tendency is quite pronounced. After betting on left and losing, for instance, 68% of *S*'s choices on the next trial are of the left key—a considerable shift up from the overall 48% of left choices. In contrast, choices of the left key drop to 28% after an unsuccessful previous choice of

[5] Frederick Mosteller and J. W. Tukey, The uses and usefulness of binomial probability paper, *J. Amer. Stat. Assn.*, 44, 1949, 174-212.

the right key. This tendency to repeat unsuccessful choices differs sharply from what happens after a win on the previous choice. Here the incidence of left and right choices is neatly in line with the overall incidence of left and right choices.

In the 70:30 gambling group, repetition after loss holds for choices of the right key (30% key) but not for the left key (70% key). The same is true in the 90:10 gambling group—repetition after loss is found for the 10% key but not for the 90% key. As with the 50:50 condition, the 70:30 and the 90:10 problem-solving groups show very little shift away from the overall proportions of left and right choices.

One interesting side effect is found in the 70:30 gambling group. This is the tendency to repeat a choice of right (30%) key after a previously successful choice of right (the incidence of right choices on such trials is 11% above the overall percentage). It appears that this effect is the result of the particular experimental schedule where two consecutive rights paid off fairly early in the series. A number of *S*s expressed surprise at this unusual event and stated that often they repeated a choice of right to determine whether or not two consecutive rights was a feature of the pattern they were hoping to find.

It is suggested that all the repetition-after-loss effects can be called negative recency-effects of the kind where Event B is anticipated following a run of Event A's. In the words of one *S* in the 50:50 group:

> It seemed to run in series of two or three each—like two lefts, then three rights, then back to two lefts and so on. I would stick to the left key for a while until I felt that the run was playing itself out. Then I'd switch to right. If right paid off, I might stick there. It depended on whether I was expecting just one right or a run of a couple of rights. If right didn't pay off, I'd stick to it. All it would mean was that the run on left was even longer than I had expected and it would be getting so long that I'd feel even surer that the right key was coming up on the next turn.

Such a use of runs as a basis for decision will cover all the tendencies to repetition after loss when supplemented by the point made by Whitfield and Smith to the effect that *S*s expect shorter runs than are found on a chance basis.[6] For the 50:50 group, *S*s apparently work on the basis of short runs of both left and right payoffs and so repetition after loss appears after choices of both keys. In the 70:30 and 90:10 groups, where a payoff on left has the higher probability of occurring, *S*s apparently try to judge the length of a run of payoffs on left and accordingly show repetition after loss on right.[7]

---

[6] J. W. Whitfield, The imaginary questionnaire, *Quart. J. Exper. Psychol.*, 2, 1950, 76-88; M. H. Smith, Spread of effect as the spurious result of non-random response tendencies, *J. Exper. Psychol.*, 39, 1949, 355-368. Where relative frequencies are equated within blocks of 10 trials, *S*s show a certain wisdom in expecting short runs. The runs expected, however, seem to be even shorter than those found in the experimental schedule. This statement is made on the basis of an analysis of runs expected by 64 *S*s in drawing 7 red and 3 yellow balls from an urn without replacement. Probably what happens is that *S*s expect short runs of 2 or 3 similar events, find that most runs are of this length, and proceed to operate on 2 or 3 as the most probable length of any run.

[7] The word "apparently" is used with advisement. Repetition after loss can have a different basis. In a 50:50 case, for example, it can be brought about by *S*'s keeping count of the number of times left and right have paid off and then betting on the key which up to then has paid off the lesser number of times. The statements of *S*s, however, indicate that judgments about length of run are the primary causes of repetition-after-loss in the present experiment.

DISCUSSION

The results of the present analysis indicate that we need not be concerned about either positive or negative recency-effects in a task where the stimulus-materials vary and the task is seen by *S*s as one not involving chance, but that we can anticipate negative recency-effects in a chance task where stimuli differ only in temporal position in a series. These results can be drawn together with other data on varieties of response-patterning in two-choice situations by considering the particular kind of response-patterning as a function of the bases for choice provided by the task and *S*'s experience. This is in some respects similar to the arguments put forward by Senders and Sowards.[8]

Where the task is seen by *S* as one involving chance and where the "stimulus on a given trial differs from the stimulus on any other trial only in its temporal position in the series," *S*s tend to use guessing habits and short sequences or runs of events as bases for choice.[9] Under such conditions, response-patterns take the form required by recency-effects. The precise form of these effects will depend on *S*'s perception of the sequential patterning of event-occurrence, and on *S*'s notions of how best to fit his responses to this patterning.

Where the task is seen by *S* as one not involving chance, and where the stimulus-material takes a number of different but recurring forms, *S*s tend to use as one basis for choice the answer they recall as having been correct for the previous presentation of a particular form of the stimulus-material.[10] In the problem-solving task discussed in the present analysis it has been found that *S*s try to remember and reproduce the answer previously correct for a particular geometrical design, which may have been presented from 2 to 19 trials before.

In the words of one *S* in the 70:30 problem-solving group, the procedure is described as follows.

> I noticed that the first cards [key cards] were the same each time around [each block of 10 trials]. So I'd try and remember the answer that had been right the time before. Sometimes I could remember for two times back and if it had been the same answer for those two times, I'd maybe change the answer when the third time came around. That didn't happen too often and I'd often get confused because I was trying to watch what the other two cards [key-cards] were like as well. Sometimes I'd go more by them than by the first card and try to remember when I'd seen cards like them before and what the answer had been. Sometimes I'd just get a feeling

---

[8] Senders and Sowards, *op. cit.*, 358-359. Senders, Further analysis of response sequences in the setting of a psycho-physical experiment, this JOURNAL, 66, 1953, 215-228.

[9] Hake and Hyman, *op. cit.*, 64.

[10] J. J. Goodnow and Leo Postman, Learning in a two-choice probability situation with a problem-solving setting, *J. Exper. Psychol.*, 49, 1955, 16-22.

about the right answer and not have to use memory at all. I'd say I was going by the first card about half the time.

Judging from *Ss*' statements and from an analysis of similarities in *Ss*' answers for the same key-cards from one block to the next, the use of memory in the problem-solving task is not as consistent as the use of runs in the gambling task. Response patterns are not as pronounced although they do occur at a statistically significant level in three of the five groups tested—in the 50:50, 80:20 and 90:10 groups but not in the 60:40 or 70:30 groups.[11] This patterning, however, is not between one trial and the immediately preceding trial but between one presentation of a form of the stimulus material and the previous presentation of the same form of stimulus material.

Where there is no knowledge of which event actually occurred, as in Senders' and Sowards' task of discriminating under sub-threshold conditions, *Ss*' own past responses provide the basis for decision. *Ss* apparently fall back on the easiest thing, repeating their previous choices until pulled up by the feeling that they have been saying the same thing for some time, and then switching to the other alternative and repeating it several times.[12] The response-patterns here take the form of relatively long runs of identical answers. It seems likely that *Ss*' judgments about when to switch would be affected by their knowledge or feeling about how frequently each event would occur, and by any attempt to keep their response-distributions in line with the expected frequencies of occurrence.

Summary

The hypothesis tested was that recency-effects in response-sequences differ with the particular stimulus-material and the way the problem is presented to *S*. Records of sequential choices were obtained for a pair of matched two-choice tasks—a gambling and a problem-solving task—and for event-probabilities of 50:50, 70:30, and 90:10. The former was a chance task and stimuli differed from one another only in temporal position in the series; negative recency-effects were consistently found. The latter was not a chance task and the stimulus material took a number of different forms; recency-effects were not found. The difference is explained in terms of the differential bases for decision provided by the two tasks.

[11] Goodnow and Postman, *op. cit.*, 20.
[12] Senders and Sowards, *op. cit.*, 371.

A SHORT METHOD FOR CALCULATING THE ADAPTATION-LEVEL FOR ABSOLUTE AND COMPARATIVE RATING JUDGMENTS

By HARRY HELSON and PHILIP HIMELSTEIN, University of Texas

Extension of the theory of adaptation-level (*AL*) to include comparative as well as absolute types of judgments has enlarged the sphere of use of equations previously published by Helson[1] and by Michels and Helson[2] for quantitatively evaluating factors which influence psychophysical and other types of judgment. Calculation of the best value of *AL* and of the weighting constant, *b*, which measures the relative contributions of the stimulus-series and of the *AL* in determining responses to stimuli, can be shortened considerably when using the method of least squares through the use of tabulated values of the function $Y = (0.5K+J) / (1.5K-J)$ which enters into Helson's equation.[3] In this article two examples are presented illustrating use of tabulated values of Y as a function of the judgments (Table I) and of coded stimulus-values (X') to minimize the computations in solving the normal equations for the desired constants. Data obtained by the method of 'absolute judgment' and by the method of comparative ratings[4] have been chosen for the worked examples in order to demonstrate application of the equation to two different scales of categories involving two psychophysical methods.

To obtain the best values of A' and b from normal equations for linear functions it is necessary to transform Helson's original equation:

$$J = K(1.5X - 0.5A' + 1.5bA')/(X + A' + bA') \quad \ldots\ldots\ldots [1]$$

into linear form as follows:

$$(0.5K + J)/(1.5K - J) = X/A' + b \quad \ldots\ldots\ldots\ldots\ldots [2]$$

where J is the average judgment of each stimulus on a numerical scale whose topmost value is K, X refers to the stimulus values, and A' and b are constants such that:

$$A = A' - bA' \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [3]$$

where A is the adaptation-level, and b is a weighting constant showing the relative contributions of the stimulus-series and of the *AL* in determining responses to stimuli. Any qualitative scale of categories may be used in judging the stimuli

* Accepted for publication November 24, 1954.

[1] Harry Helson, Adaptation-level as a basis for a quantitative theory of frames of reference, *Psychol. Rev.*, 55, 1948, 297-313.

[2] W. C. Michels and Harry Helson, A reformulation of the Fechner law in terms of adaptation-level applied to rating-scale data, this JOURNAL, 62, 1949, 355-368.

[3] Helson, *op. cit.*, 304-305.

[4] Harry Helson, W. C. Michels, and A. Sturgeon. The use of comparative rating scales for the evaluation of psychophysical data, this JOURNAL, 67, 1954, 321-326.

provided that there are $2N + 1$ categories in the scale of which N categories are above and N are below the middle category. Similarly, any numerical scale may be employed into which the qualitative categories may be translated for computational purposes provided that: (1) the numerical scale is linear with the qualitative scale; and (2) the topmost value, K, of the numerical scale is twice the value assigned to the middle category. All numerical scales satisfying these two criteria will yield identical values for any given scale of categories since the left-hand side of Equation [2] is a ratio which is a linear function of the stimulus-values.[5]

It is obvious that Equation [2] is a linear function of the form $Y = mX + b$ with the left-hand side equal to Y, $1/A'$ is equal to the slope, m, and b is the y-intercept. A somewhat simpler expression for Y is obtained if the J's are taken from the center of the numerical scale so that $J' = J - 0.5K$ giving:

$$(K + J') \,/\, (K - J') = (0.5K + J) \,/\, (1.5K - J) \quad \ldots\ldots\ldots\ldots [4]$$

Table I contains values of Y as a function of J on a numerical scale from 0.00 to 9.95 ($K = 10$) in steps of 0.05. By means of this table Y is given immediately from J.

Equation [1] may be written in terms of A rather than A' by substituting $A / (1 - b)$ for A'—from Equation [3]—so as to give:

$$J = K(X - A) \,/\, [X + (1 + b) \,/\, (1 - b)A] + 0.5K \quad \ldots\ldots\ldots [5]$$

After determining A' and b by means of normal Equations [6] and [7], and A by means of Equation [3], theoretical values of J may be calculated either by means of Equation [1] if the value of A' is used, or by means of Equation [5] if the value of A is employed.

From Equation [5] it is seen that b determines the relative contributions of the adaptation-level, A, and the stimuli, X, to the judgments since $(1 + b) \,/\, (1 - b)$ is a weighting constant. As shown in a recent study by Helson, Michels, and Sturgeon,[6] when a standard is employed and the judgments are in terms of a comparison between variables and standard, the adaptation-level is largely determined by the standard hence the weighting constant for A reflects the contribution of the standard to the judgmental process.

Before turning to examples of the computations, let us summarize the information gained by obtaining the best-fitting values of A and b from a set of observed data. We find first that the theoretical J-values represent psychological intervals on a true interval-scale. Our procedure results in scaling the stimuli on a psychological continuum. Secondly, we obtain the best value of the adaptation-level from the observed data by method of least squares. Thirdly, we can determine quantitatively the relative contributions of the stimulus-series and of the adaptation-level in the judgments. The weighting factor for the adaptation-level includes everything in the judging situation which influences the results apart from the stimulus-series whose weighting factor is unity. If a standard is employed, as in the comparative rating-scale, the value of the adaptation-level is largely determined by the standard, and the weighting constant $(1 + b) \,/\, (1 - b)$ gives a good indication of the extent to which the standard determines the judgment as compared with the stimulus-series.[7]

---

[5] For proof of the sufficiency of these conditions and the theoretical derivation of the original equation, see Helson, *op. cit.*, 302ff.

[6] Helson, Michels, and Sturgeon, *op. cit.*, 321 ff.

[7] For more detailed discussion of this point, see Helson, Michels, and Sturgeon, *op. cit.*, 324ff.

Examples: (1) Data obtained by method of absolute judgment. The experimental data in Table II, represented by the J-values in column 2, were obtained by the method of single stimuli (absolute judgment) with no comparison or background stimulus.

The stimuli consisted of 5 weights ranging from 200 to 400 gm. in steps of 50 gm. and were presented in random order. Two series of observations were made for practice, followed by five experimental series. Every J-value is the average of 20 observations (4 *O*s $\times$ 5 series). The category scale used in judging the stimuli and the numbers assigned to the categories were as follows:

very, very heavy	9
very heavy	8
heavy	7
medium heavy	6
medium	5
medium light	4
light	3
very light	2
very, very light	1

From the numbers assigned the qualitative categories it is evident that we are using a 0–10 numerical scale with 5 for the neutral judgment and 10 for the topmost value of the scale. Categories for extreme values of the scale, 0 and 10, are not given because experience has shown that if the ends of the scale are left open and *O*s are allowed to introduce categories of their own choosing for the extremes, better results are obtained in that the *O*s do not avoid the end categories actually provided thus reducing central tendency effect to some extent. Further, it has been found that the effect of changing stimulus-series or of introducing background stimuli is reflected better in the judgments when *O*s judge in terms of qualitative categories than when judging in terms of numerical scales. If *O*s are given numerical scales, then it is necessary to define at least the middle and end points qualitatively for satisfactory use of numbers in judging stimuli.

To expedite computations, Table I has been prepared giving values of the function $Y = (0.5K+J) \,/\, (1.5K-J)$. Its values have been tabulated for a J-scale running from 0 to 10 in steps of 0.05. By entering Table I with observed values of J the value of Y is found immediately. By coding the X-values the least squares solution for $1/A'$ is considerably simplified. Letting X' stand for coded values of X, the normal equation for the solution of $1/A'$ is:

$$1/A' = (\Sigma X'Y - nM_yM_x') \,/\, (\Sigma X'^2 - nM^2_{x'}) \quad \ldots\ldots\ldots [6]$$

and for b is:

$$b = M_y - (1/A')M_x \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [7]$$

By coding X so that $M_{x'}$ equals 0, the second term in both the numerator and denominator of Equation [6] drops out and the coded squares can be immediately written down without the use of tables. The value yielded by Equation [6] for $1/A'$ must, however, be divided by the step-interval

TABLE I

Tabulated Values of $Y=(0.5K+J)/(1.5K-J)$ as a Function of J

J	Y	J	Y	J	Y	J	Y
0.00	0.3333	2.50	0.6000	5.00	1.0000	7.50	1.6667
0.05	0.3378	2.55	0.6064	5.05	1.0101	7.55	1.6846
0.10	0.3423	2.60	0.6129	5.10	1.0202	7.60	1.7027
0.15	0.3468	2.65	0.6194	5.15	1.0305	7.65	1.7211
0.20	0.3514	2.70	0.6260	5.20	1.0408	7.70	1.7397
0.25	0.3559	2.75	0.6327	5.25	1.0513	7.75	1.7586
0.30	0.3605	2.80	0.6393	5.30	1.0619	7.80	1.7778
0.35	0.3652	2.85	0.6461	5.35	1.0725	7.85	1.7972
0.40	0.3699	2.90	0.6529	5.40	1.0833	7.90	1.8169
0.45	0.3746	2.95	0.6598	5.45	1.0942	7.95	1.8369
0.50	0.3793	3.00	0.6667	5.50	1.1053	8.00	1.8571
0.55	0.3841	3.05	0.6736	5.55	1.1164	8.05	1.8777
0.60	0.3889	3.10	0.6807	5.60	1.1276	8.10	1.8986
0.65	0.3937	3.15	0.6878	5.65	1.1390	8.15	1.9197
0.70	0.3986	3.20	0.6949	5.70	1.1505	8.20	1.9412
0.75	0.4035	3.25	0.7021	5.75	1.1622	8.25	1.9629
0.80	0.4085	3.30	0.7094	5.80	1.1739	8.30	1.9851
0.85	0.4134	3.35	0.7167	5.85	1.1858	8.35	2.0075
0.90	0.4185	3.40	0.7241	5.90	1.1978	8.40	2.0303
0.95	0.4235	3.45	0.7316	5.95	1.2099	8.45	2.0534
1.00	0.4286	3.50	0.7391	6.00	1.2222	8.50	2.0796
1.05	0.4337	3.55	0.7467	6.05	1.2346	8.55	2.1008
1.10	0.4388	3.60	0.7544	6.10	1.2472	8.60	2.1250
1.15	0.4440	3.65	0.7621	6.15	1.2599	8.65	2.1496
1.20	0.4493	3.70	0.7699	6.20	1.2727	8.70	2.1746
1.25	0.4345	3.75	0.7778	6.25	1.2857	8.75	2.2000
1.30	0.4599	3.80	0.7857	6.30	1.2989	8.80	2.2258
1.35	0.4652	3.85	0.7937	6.35	1.3121	8.85	2.2520
1.40	0.4706	3.90	0.8018	6.40	1.3256	8.90	2.2787
1.45	0.4760	3.95	0.8099	6.45	1.3392	8.95	2.3058
1.50	0.4815	4.00	0.8182	6.50	1.3529	9.00	2.3333
1.55	0.4870	4.05	0.8265	6.55	1.3669	9.05	2.3613
1.60	0.4925	4.10	0.8349	6.60	1.3809	9.10	2.3898
1.65	0.4981	4.15	0.8433	6.65	1.3952	9.15	2.4188
1.70	0.5038	4.20	0.8518	6.70	1.4096	9.20	2.4483
1.75	0.5094	4.25	0.8605	6.75	1.4242	9.25	2.4783
1.80	0.5152	4.30	0.8692	6.80	1.4390	9.30	2.5088
1.85	0.5209	4.35	0.8779	6.85	1.4540	9.35	2.5398
1.90	0.5267	4.40	0.8868	6.90	1.4691	9.40	2.5714
1.95	0.5326	4.45	0.8957	6.95	1.4845	9.45	2.6036
2.00	0.5385	4.50	0.9048	7.00	1.5000	9.50	2.6364
2.05	0.5444	4.55	0.9139	7.05	1.5157	9.55	2.6697
2.10	0.5504	4.60	0.9231	7.10	1.5316	9.60	2.7037
2.15	0.5564	4.65	0.9324	7.15	1.5478	9.65	2.7383
2.20	0.5625	4.70	0.9417	7.20	1.5641	9.70	2.7736
2.25	0.5686	4.75	0.9512	7.25	1.5806	9.75	2.8095
2.30	0.5748	4.80	0.9608	7.30	1.5974	9.80	2.8462
2.35	0.5810	4.85	0.9704	7.35	1.6[illegible]44	9.85	2.8835
2.40	0.5873	4.90	0.9802	7.40	1.6316	9.90	2.9216
2.45	0.5936	4.95	0.9900	7.45	1.6490	9.95	2.9604

between stimuli in order to yield correct values of A' and A. It should also be noted that the mean of the original X-values must be used in Equation [7] for determination of b.

The computations leading to A and b and the theoretical J-values are given in Table II. It is seen that the fit to the observed data is excellent and that the stimulus-series exerts about three times as much weight as the

TABLE II

DERIVATION OF THE CONSTANTS A' AND b IN EQUATION [2]

(Data obtained by method of absolute judgment. From Helson, *op. cit.*, *Psychol. Rev.*, 301)

X	J	Y (From Table I)	X' (Coded X)	$X'Y$	X'^2	J_t
200	3.4	0.7241	−2	−1.4482	4	3.3
250	5.0	1.0000	−1	−1.0000	1	5.1
300	6.3	1.2989	0	0.0000	0	6.4
350	7.4	1.6316	1	1.6316	1	7.4
400	8.2	1.9412	2	3.8824	4	8.2
		$M_y=1.31916$	$M_x'=0$	$\Sigma X'Y=3.0658$	$\Sigma X'^2=10$	

$$1/A'=(3.0658/10)/50=0.0061316$$
$$b=1.31916-(0.0061316\times 300)=-0.52032$$
$$A=A'-bA'=163.09-(-84.86)=247.95$$
$$J_t=10(X-A)/[X+(1+b)/(1-b)A]+0.5K$$
$$=10(X-247.95)/(X+78.24)+5.0$$

adaptation-level in determining the judgments of the stimuli, a finding in accordance with earlier work with lifted-weights reported by one of the authors.[8]

(2) *Data obtained by method of comparative ratings.* Our second example involves use of the comparative rating-scale introduced recently by Helson, Michels, and Sturgeon.[9] In this method a standard is employed and Os judge the variables with respect to standard in terms of comparative categories.

The data in Table III represent means of 5 series of judgments of 5 stimulus-weights judged by 12 Os according to the following categories:

very much heavier	9
much heavier	8
heavier	7
a little heavier	6
equal	5
a little lighter	4
lighter	3
much lighter	2
very much lighter	1

[8] Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29.

[9] *Op. cit.*, 321-326.

The numbers opposite the categories were assigned, as in the case of the method of absolute judgment, for computational purposes on a 0–10 numerical scale. Extreme categories were again purposely omitted.

TABLE III

COMPUTATION OF THE ADAPTATION-LEVEL AND SCALE VALUES OF LIFTED-WEIGHT STIMULI FOR DATA OBTAINED BY THE METHOD OF COMPARATIVE RATINGS
(Data from Helson, Michels, and Sturgeon, *op. cit.*, 322)

X	J	Y (From Table I)	X' (Coded X)	$X'Y$	X'^2	J_t
300	3.3	0.7094	−2	−1.4188	4	3.1
350	4.3	0.8692	−1	−0.8692	1	4.4
400	5.3	1.0619	0	0.0000	0	5.4
450	6.3	1.2989	1	1.2989	1	6.3
500	7.1	1.5316	2	3.0632	4	7.0
		$M_y=1.0942$	$M_x'=0$	$\Sigma X'Y=2.0741$	$\Sigma X'^2=10$	

$$1/A'=(2.0741)/(10)=0.20741/50=0.0041482$$
$$b=1.0942-(0.0041482\times 400)=-0.5651$$
$$A=A'-bA'=241.07-(-136.25)=377.32$$
$$J_t=10(X-A)/[X+(1+b)/(1-b)A]+-0.5K$$
$$=10(X-377.32)/(X+104.82)+5.0$$

In this method, since the middle category corresponds to 'Equal,' the value of the adaptation-level, A, is the value of stimulus eliciting the equality judgment and it is therefore comparable to the point of subjective

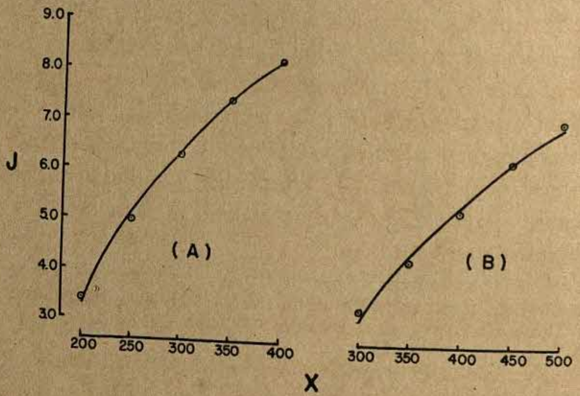


FIG. 1. CURVES OF THEORETICAL AND OBSERVED VALUES
The theoretical values are shown by the solid lines; the observed values by the small circles. Curve (A) illustrates the results obtained by the method of absolute judgment: Curve (B) by the method of comparative ratings).

equality (*PSE*). The computed value of A, 377.32, shows that there is a negative time-order error with standard-first, variable-second, as the value of the standard is 400 gms. Since the negative time-order effect is not

large and shows up clearly with both the absolute and comparative rating-scale methods, as we have employed them, we can say that they are capable of yielding measures of discrimination as fine as the longer, classical procedures.

Graphs of Equation [1] are given in Fig. 1 which are plotted according to the constants derived from the data in Tables II and III. From these plots it is seen that Equation [1] yields a negatively accelerated curve closely approximating a logarithmic function. In fact Equation [1] and the logarithmic equation of Michels and Helson yield approximately identical values of A and equally good fits to a variety of data.[10] Since the function defined by Equation [1] is negatively accelerated, larger and larger increments of stimulus are needed for successive equal differences in perception and judgment. The excellence of the fits to observed data (plotted circles in Fig. 1) shows that the theory underlying Equation [1] is adequate for data obtained both by method of absolute judgment and by comparative rating procedures employing standards for comparison with series stimuli.

In conclusion it should be pointed out that when the methods described here for obtaining data and Equation [1] are employed to determine adaptation-level and to scale stimuli, there is considerable saving in time and labor in both the collection of the data and in the computations as compared with classical procedures. We have found in many experiments with a large number of *O*s that five experimental series following two practice series are sufficient to obtain reliable data with absolute or comparative judgments as contrasted with the 50 or 100 series recommended for the methods of constant stimuli and constant stimulus-differences. Whereas the older methods require different equations to determine *PSE* and threshold-values, Equation [1] yields both the value of adaptation-level and the judgmental scaled values of the stimuli in one set of computations. In addition, Equation [1] furnishes a unified theory of absolute and comparative types of judgments as well as a theory of a number of phenomena associated with them.[11]

---

[10] The best value of A according to the Michels-Helson logarithmic equation for the data of Table III is 381 gm. as against 377 gm. by Equation [1]—a difference of about 1%.

[11] For example, time-order errors as shown by Michels and Helson, A quantitative theory of time-order effects, this JOURNAL, 67, 1954, 327-334.

THE EFFECT OF ELECTRO-SHOCK THERAPY ON THE RATES OF REVERSAL OF AMBIGUOUS PERSPECTIVE FIGURES

By V. R. Fisichelli, Hunter College, and F. V. Rockwell and Lenore Clarke, Cornell University Medical College

It has been demonstrated in several studies that electro-shock therapy affects the peripheral visual apparatus in a number of ways.[1] It has further been demonstrated that it is responsible, directly or indirectly, for changes that are mediated by higher levels of function or that are primarily psychological in nature.[2] The present study was undertaken in the hope of shedding additional light on the problem of the relationship between electro-shock therapy and such visual processes. The experiment was performed in the further hope that some information might be gathered in respect to the reversibility of ambiguous figures in depressed patients.

Procedure

Two groups of depressed patients, each containing 20 members, were given two test-sessions in which their rates of reversal were measured. The two test-sessions were separated by an interval of 2 wk. during which time the Experimental Group of patients was given four electro-shock treatments while the Control Group of patients received none. The retest session took place exactly 24 hr. after the last shock treatment. Each patient in the shock group received charges of 60 cps. alternating current at a sufficient voltage and for the length of time necessary to produce a grand-mal convulsion at each treatment. The charges varied from 90 to 130 v. and the lengths of time from 0.1 to 0.5 sec.

The patients in the Experimental and Control Groups were equated as nearly as possible with respect to age, sex, and type of mental disorder. While mixed diagnoses were found in each group with approximately the same frequency, all patients were depressed.

The material for the test of reversibility consisted of a 16-mm. film on which 10 test-figures were shown in a series. Two different ambiguous perspective diagrams were used, each shown 5 times in consecutive runs of 1-min. duration. The runs were separated by 10 sec. of blank film. The two figures used were the conventional

* Accepted for publication July 26, 1954. The authors are grateful to Drs. O. Diethelm and R. Kohl of Payne Whitney Clinic for their coöperatioon and assistance.

[1] W. Karliner, Neurologic signs and complications of electric shock treatments, *J. Nerv. & Ment. Dis.*, 107, 1948, 1-10. J. McClellan, M. Morriss, and J. Marlson, Measurement of intra-ocular tension in insulin and electroconvulsive therapy, *Bull. Menninger Clinic*, 14, 1950, 220-226.

[2] O. Diethelm and P. Regan, Postconvulsive psychopathological states in electric convulsive treatment, *Psychosomat. Med.*, 12, 1950, 78-85.

static book-figure and the animated Lissajous figure described in previous studies by one of the present authors.[3]

The film was projected with a 750-w. Bell and Howell projector placed in one room with *S* in an adjacent room, viewing the film through an opaque glass panel, 10 × 14 in. in size. The projected image was approximately 6 × 6 in. The projector was controlled remotely by *E* who sat next to and slightly behind *S*. *S*'s room was dimly illuminated.

Each *S* was seated before the screen with instructions to watch the first run of the figure and report what he saw. Once he reported reversals spontaneously he was further instructed as follows:

I'm going to show you this figure again several times. Your task will be to watch it as you just did and to tell me at all times precisely what you see. You must be especially careful to tell me when it changes and how it has changed. If you have any lengthy remarks or comments to make it would be best to save them for the rest-periods between figures; otherwise you may go right on talking providing you do not forget to tell me when the changes occur.

After the practice run and the instructions were given, *S* was advised that this was not a test of any special ability, that he would not be scored on the number of reversals seen, and that he might see many or few, if any, reversals. He was finally instructed to watch the figure as a whole, not fixating any point or points, and to avoid trying to produce reversals or to suppress them. Half of the *S*s were shown the book-figure first and half were shown the Lissajous figure first.

Results

The total number of reversals seen by each patient for the separate figures of the test- and retest-sessions are presented in Table I. The numbers appearing in the body of the table represent sums for four trials of 1 min. each. It will be seen immediately that the mean rate of reversal for the patients who received the shock treatment drops to about half its original size while that for the Control Group of patients remains about the same. The decrement is evidenced for both the static book-figure and the moving Lissajous figure.

When the differences between test and retest in mean rate of reversal were subjected to a *t*-test, values were obtained which indicate that the decrease in rate of reversals is statistically significant for the Experimental Group receiving shock. The differences found for the Control Group are not statistically significant. The differences and the *t*-values obtained for them are presented in Table II.

Further examination of Table I reveals rather marked individual differences among both groups of patients. It is interesting to note, however,

[3] V. R. Fisichelli, An experimental study of some factors involved in the fluctuations of ambiguous perspective figures, unpublished Ph.D. dissertation, Fordham University, 1948. Fisichelli, Cinematic Lissajous figures: their production and applicability in experiments on reversible perspective, this JOURNAL, 64, 1951, 402-405.

TABLE I

TOTAL NUMBER OF REVERSALS SEEN BY EACH PATIENT FOR THE SEPARATE FIGURES OF THE TEST- AND RETEST-SESSIONS

IM=involutional melancholia; *Ne*=neurotic; *CP*=constitutional psychopath; *Sc*=schizophrenia; *NG*=not given. *Ss* 1–6 were men; 7–20 were women.

	Subject	Diagnosis	Test			Retest		
			Liss.	Book	Both	Liss.	Book	Both
Experimental Group	1	*IM*	11	1	12	5	2	7
	2	*Ne*	0	11	11	2	3	5
	3	*NG*	1	1	2	0	0	0
	4	*NG*	3	0	3	0	5	5
	5	*NG*	5	17	22	2	9	11
	6	*NG*	16	42	58	11	9	20
	7	*CP*	4	5	9	0	0	0
	8	*Sc*	15	0	15	1	0	1
	9	*Sc*	0	1	1	1	1	2
	10	*Sc*	5	0	5	10	0	10
	11	*Sc*	10	13	23	0	0	0
	12	*Sc*	2	0	2	3	3	6
	13	*Sc*	21	32	53	2	13	15
	14	*IM*	27	6	33	5	2	7
	15	*IM*	2	10	12	1	8	9
	16	*Ne*	4	2	6	6	2	8
	17	*Ne*	5	0	5	8	0	8
	18	*Ne*	15	44	59	10	28	38
	19	*NG*	43	36	79	15	8	23
	20	*NG*	31	9	40	14	14	28
Total			220	230	450	96	107	203
Mean			11.0	11.5	22.5	4.8	5.4	10.2
SD			11.4	14.5	22.7	4.8	6.8	9.8
Control Group	1	*CP*	15	0	15	16	4	20
	2	*Sc*	4	6	10	12	3	15
	3	*Sc*	23	59	82	56	74	130
	4	*Sc*	4	9	13	2	0	2
	5	*Ne*	5	22	27	2	0	2
	6	*NG*	38	0	38	43	1	44
	7	*Sc*	11	13	24	14	10	24
	8	*Sc*	26	17	43	30	15	45
	9	*Sc*	3	8	11	8	3	11
	10	*Sc*	15	1	16	20	7	27
	11	*IM*	64	14	78	57	26	83
	12	*IM*	73	0	73	76	7	83
	13	*IM*	35	4	39	0	0	0
	14	*Ne*	2	40	42	1	41	42
	15	*Ne*	11	6	17	10	1	11
	16	*Ne*	17	63	80	13	28	41
	17	*Ne*	13	20	33	18	34	52
	18	*Ne*	22	4	26	12	11	23
	19	*Ne*	30	56	86	55	71	126
	20	*Ne*	7	12	19	10	16	26
Total			418	354	772	455	352	807
Mean			20.9	17.7	38.6	22.8	17.6	40.4
SD			19.0	19.8	25.8	21.8	21.8	37.1

that while the overall variability of the Experimental Group drops sharply after four electro-shock treatments it remains about the same in the Control Group. It is evident, therefore, that individual differences are diminished in that group of patients which received electro-shock therapy.

TABLE II

THE DIFFERENCES BETWEEN MEAN RATES OF REVERSAL ON TEST AND RETEST, AND t-VALUES FOR EACH DIFFERENCE

Group	Figures	Diff.	t
Experimental	Liss.	− 6.20	3.01*
	Book	− 6.15	2.59†
	Both	−12.35	3.31*
Control	Liss.	1.85	.63
	Book	− .10	.04
	Both	1.75	.37

* Significant at 1-% level. † Significant at 2-% level.

A significant feature of the results given in Table I concerns the difference in the mean rate of reversal between the Experimental Group and the Control Group of patients on the initial test. For both figures combined this difference exceeds the 5-% level of confidence. It cannot be readily accounted for by any variable

TABLE III

MEAN RATE OF REVERSAL FOR EACH TRIAL ON THE TEST AND RETEST ACCORDING TO FIGURE PRESENTED AND GROUP MEASURED

	Experimental Group		Control Group	
Figure	Test	Retest	Test	Retest
Lissajous	Practice run		Practice run	
	2.30	.80	3.25	3.90
	2.60	1.00	5.00	5.40
	3.00	1.30	5.95	6.85
	3.10	1.70	6.70	6.60
Total	11.00	4.80	20.90	22.75
Book	Practice run		Practice run	
	1.05	.90	3.30	3.35
	2.50	1.20	4.25	5.70
	3.20	1.40	4.90	3.65
	4.75	1.85	5.25	4.90
Total	11.50	5.35	17.70	17.60

known to the present experimenters, although there is a slight suggestion that a difference in age might account for some of the difference.

In Table III the mean rates of reversal are presented for each trial on the test- and retest-sessions according to the figure shown and the group measured. The major points of difference between the two groups of patients will be seen again. These data emphasized, however, the consistency of those differences from trial to

trial or from minute to minute. The slight increments from minute to minute that are indicated for both groups on the initial tests with each figure were not subjected to statistical analysis. They are consistent, however, with known effects of duration of exposure.[4]

Test-retest product-moment correlation coefficients for the two groups of patients are presented in Table IV. The interval between test and retest was 2 wk., and the product-moment computations were based on those values already presented in Table I.

It is evident immediately that the coefficients of reliability are uniformly higher for the Control Group. All of the coefficients are significant beyond the 1-% level of confidence.

TABLE IV

COEFFICIENTS OF TEST-RETEST CORRELATION

Figure	Experimental Group	Control Group
Lissajous	.66	.81
Book	.76	.84
Both	.78	.85

DISCUSSION

The most striking feature of the present findings is the sharp decrement in reversal rate for the Experimental Group after four electro-shock treatments. This decrement in rate of reversal correlates with two outstanding changes in the *Ss* following electro-shock therapy: the first is clinical improvement in mood and the second is the presence of a certain amount of mild and reversible brain damage. Unfortunately, the nature of the present experimental design makes it difficult to determine whether the decrement is mediated by a primary organic change or by a secondary affective one. It is strikingly clear, however, that quite apart from the ultimate mechanism involved, electro-shock therapy brought about the change in reversal rate. A replication of the present study using insulin or some other method of shock might be helpful in clarifying the explanation.

The fact that the rate of reversal for the control patients showed no significant change after 2 wk. is interesting not only because it provides a standard of comparison with the experimental *Ss* of this study, but also because it is at variance with results obtained from normal *Ss*. When 12 women students from Hunter College were subjected to tests identical with those described in this study, their mean rate of reversal showed a statistically significant increase on the retest-session two weeks later. The testing conditions for these students were identical with those of the Control Group, and while these students were used principally to check on the comparability of the present testing situation with other situations, the results

[4] Wolfgang Köhler, *Dynamics in Psychology*, 1940, 71. Fisichelli, Reversible perspective in Lissajous figures: Some theoretical considerations, this JOURNAL, 60, 1947, 241.

obtained from them shed additional light on the problem under consideration. Their mean rates of reversal and standard deviations are given in Table V. These values may be compared directly with the means and standard deviations found in Table I above.

When the increase in rate of reversal on the retest over the test was analyzed for statistical significance the following *t*-ratios were obtained: Lissajous figure, 1.89, book-figure, 1.78, both, 2.27. The last of these is significant at the 5-% level of confidence. In all probability the increase in the student rate of reversal is further evidence of some kind of a practice-effect; but, in any case, it

TABLE V

RATES OF REVERSAL FOR A GROUP OF STUDENTS ON TEST AND RETEST UNDER CONDITIONS COMPARABLE WITH THOSE USED WITH CONTROL GROUP

	Test			Retest		
	Liss.	Book	Both	Liss.	Book	Both
Mean	20.00	13.17	33.17	24.99	21.05	46.04
SD	9.70	13.73	18.06	10.40	21.47	27.16

is evident that the depressed control patients were not influenced by it. The increment found for the students also serves to emphasize by contrast the decrement seen for the patients who received shock therapy.

A recent study by Philip has as one of its principal findings that electro-shock does not influence the rate of reversal of an ambiguous figure.[5] Since the results of this study appear to conflict with those of the present paper some clarification of this difference would seem to be desirable.

In the Philip study a large group of patients were given tests of reversible perspective for the purpose of comparing the frequency of reversals for manic-depressives and schizophrenics. Further, since many of the *S*s had had previous electro-shock therapy at one time or another, the effect of the treatment on the reversal score was explored. The rates of reversal of those patients who had received some treatment with electro-shock therapy were compared with the rates of those patients who had received no electro-shock therapy. No significant difference was found between them, although there was some indication that those who received no treatment tended to score higher. A further comparison of those whose treatment was recent (within 24 hr.) with those whose treatment was remote yielded no significant difference.

In view of the rather high test-retest correlation shown in the present study, it is clear that a measurement of reversibility made before and after electro-shock therapy on the same *S*s will have a much greater chance of discovering any effect, if it does exist. In this way the rather large individual differences pointed out in the present study can be controlled. Furthermore no assumptions have to be made with respect to the homogeneity of the groups compared with regard to age, sex, diagnosis, and other possible variables. Thus the present study would appear to be better designed to reveal the differences in question.

[5] B. R. Philip, Reversals in the perception of Lissajous figures by psychotics, *Canadian J. Psychol.*, 7, 1953, 115-125.

Summary

The present study was designed to measure the effect of electro-shock therapy on the rate of reversal of figures of ambiguous perspective. Accordingly two groups of depressed patients were subjected to tests of reversibility using the moving Lissajous figure and the stationary book-figure. The experimental group of patients was tested before and after four electro-shock treatments were administered within a period of two weeks. The control group of patients was tested in the same way but received no shock therapy during the 2-wk. interim. The test results indicate that:

1. Electro-shock tends to decrease the reversal rate.
2. With respect to measures of reversibility, both groups of patients tend to maintain their positions within their respective groups even over a period of 2 wk.

AN EXPERIMENTAL COMPARISON OF SEVERAL PSYCHOLOGICAL SCALES OF WEIGHT

By RALPH R. CANTER, The Rand Corporation, and JERRY HIRSCH, University of California

A psychological scale of weight constructed by the method of fractionation is quite different from a scale constructed on the assumption that all *JND*s are subjectively equal.[1] In the former, perceived weight is represented as a positively accelerated function of physical weight—the physical increments in an equal-interval scale get progressively smaller. In the latter case, perceived weight is represented as a negatively accelerated function of physical weight—the physical increments in an equal-interval scale get progressively larger. As Harper and Stevens have noted, reliance on the method of fractionation makes it necessary to assume that *JND*s become subjectively larger as physical weight increases.[2] In the experiment to be reported, the two scaling principles were evaluated in terms of the responses of *S*s to series of weights constructed according to both principles. A linear series of weights, in which the increments were physically equal, provided a convenient control series.

Method. Four series of weights (Veg, Linear, Point-*JND,* and Spread-*JND*), whose physical values are shown in Table I, were constructed. Each of the first

TABLE I

WEIGHT SERIES (IN GM.)

Category	Veg	Linear	Point-*JND*	Spread-*JND*
1	43	43	43	40, 41, 43, 45, 47
2	98	68	52	48, 50, 52, 54, 56
3	131	93	63	59, 61, 63, 66, 68
4	159	118	77	71, 74, 77, 80, 83
5	183	143	93	86, 89, 93, 96, 100
6	204	168	112	104, 108, 112, 117, 121
7	224	193	136	126, 131, 136, 141, 147
8	243	218	165	153, 159, 165, 171, 178
9	260	243	199	185, 192, 199, 207, 215
10	277	267	241	224, 232, 241, 251, 261
11	292	292	292	271, 281, 292, 303, 316

three series contained 55 weights extending over the same range (43.18–292.38 gm.) and each was composed of 11 successively heavier groups of 5 equal weights. The fourth series consisted of 55 successively heavier weights ranging from 40.00 to 315.63 gm. The Veg series (decreasing increment) represented an equal-interval scale according to the Veg formula: $\log V = 14.58 \log (1 + \log W) - 6.94$.[3] Two

* Accepted for publication December 15, 1954.

[1] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology,* Rev. ed., 1954, 242-243. J. P. Guilford, *Psychometric Methods,* 2nd ed., 1954, 208-213.

[2] R. L. Harper and S. S. Stevens. A psychological scale of weight and a formula for its derivation, this JOURNAL, 61, 1948, 343-352.

[3] Harper and Stevens, *op. cit.*, 346.

JND series (increasing increment) represented equal-interval scales according to the Weber-Fechner law. In the Point-*JND* series the successive groups of 5 equal weights were about 5 *JND*s (19.5%) apart. In the Spread-*JND* series each of the 55 different weights was about 1 *JND* (3.9%) apart. The mean weight of each successive set of five weights equalled the weight assigned to each successive group in the Point *JND* series. The Linear series (constant increment) was physically linear, with increments of 25 gm.

Subjects. The *S*s were 40 students of introductory psychology at the University of California. One group of 20 *S*s judged the Veg and Spread-*JND* series, while the other group of 20 *S*s judged the Point-*JND* and Linear series (the order of judgment being counter-balanced within each group). Thus, the analysis of primary interest—a comparison of the Veg and Point-*JND* series—could be made in terms of independent data. A comparison of Point-*JND* and Spread-*JND* series, which were expected to give identical results, also could be made in terms of independent judgments.

Procedure. S was seated at a table, the top of which was marked off into 11 different spaces of approximately equal size and numbered from 1 to 11. Before him were 55 randomly ordered weights and a set of written instructions. His task was to distribute the weights among the 11 categories according to subjective heaviness as in Thurstone's method of equal-appearing intervals. After work with the first series of weights was completed, a second series was presented.

Results. In Fig. 1, the results for the four series are plotted in terms of the mean of the categories to which the weights in each successive set of five were assigned. Inspection of the figure shows that the two *JND* series, functionally indistinguishable as predicted, give the closest approximation to a linear plot. The Veg series gives the poorest approximation to linearity, with a fairly large number of responses being crowded into the upper categories.

Curves could be fitted to these response-distributions by the general polynomial regression equation, $y = a + bx + cx^2 + dx^3 + ex^4 + \ldots$ Anderson and Houseman describe a method which reduces the curve-fitting labor and provides tests of significance of the amount of decrease in the residual sums of squares due to fitting equations of successively higher degree by the method of least-squares.[4] The results of these tests are summarized in Table II.[5] They reveal that the question of what stimulus-distribution produces an exactly linear response-distribution still remains unanswered, since a first degree equation does not give the best fit to any of the obtained distributions. It is obvious that while the *JND*-distributions give a much closer approximation to that elusive relationship than do the Veg and Linear series,[6] they do not stand up under rigorous testing.

Discussion. Is it possible to account for the discrepancy between the present results and those obtained by Harper and Stevens by the method of fractionation? If

[4] R. L. Anderson and E. E. Houseman, Tables of orthogonal polynominal values extended to N = 104, *Iowa Agr. Exp. Sta. Res. Bull.*, 297, 1942, 1-77.

[5] The authors are indebted to Professor T. A. Jeeves and Mr. Henry Kaiser for assistance in this statistical analysis.

[6] When analyzed in terms of information theory (W. R. Garner and H. W. Hake, The amount of information in absolute judgments, *Psychol. Rev.*, 58, 1951, 446-459), transmitted information is greatest for the *JND*, second for the Linear, and least for the Veg series.

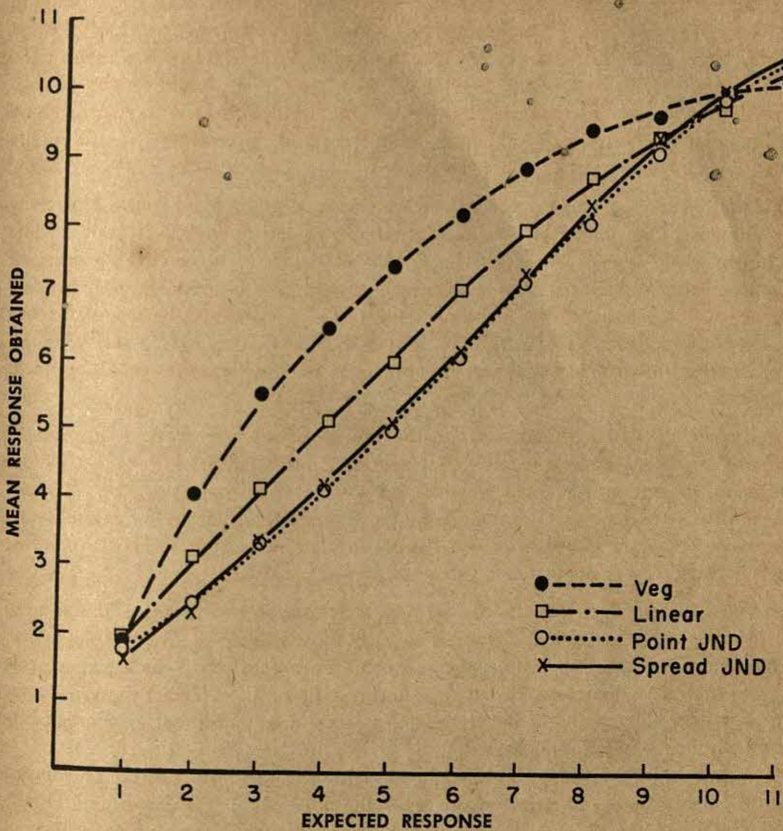


FIG. 1

TABLE II

TESTS OF SIGNIFICANCE OF REGRESSION COEFFICIENTS

Source	df.	Point-JND		Veg		Linear	
		Mean square	F	Mean square	F	Mean square	F
Linear regression	1	96.5689	3433.000*	68.2377	94.407*	82.0152	52.778*
Deviations from linear regression	9	0.0281		0.7228		0.1372	
Second-degree term	1	0.0024	—	6.1543	140.189*	1.1433	99.417*
Deviations from quadratic regression	8	0.0313		0.0439		0.0115	
Third-degree term	1	0.2000	27.780*	0.2460	16.564*	0.0617	14.349*
Deviations from third-degree regression	7	0.0072		0.0149		0.0043	
Fourth-degree term	1	—	—	0.0479	5.096†	—	—
Deviations from fourth-degree regression	6	—		0.0094		—	

* $P<.01$.
† $P<.10$.

we assume that the determination for each standard weight used by Harper and Stevens were separated in time, and were made in random order—*i.e.* that they used eight separate judgmental situations with approximately zero interaction among them—then we may expect a separate adaptation-level (*AL*) for each judgmental situation. When the level of adaptation for each set is calculated by means of Helson's second equation,[7] we find that the distribution of *AL* values fits very closely the distribution of the Harper-Stevens median-weights-judged-half (Table III). Certain discrepancies appear, particularly at the 500- and 2000-gm. levels, which may perhaps be attributed to the use of four different sizes of weight-container in the Harper-Stevens experiment—half-pint cans at the 500-gm. level and pint cans

TABLE III

Comparison of the Values Obtained by Harper-Stevens and the Adaptation-Level
(All values in grams.)

Standard	Median-weight-judged-half	Calculated *AL*
20	15.80	15.33
40	28.00	29.77
70	51.70	52.10
100	77.00	82.47
300	195.00	164.96
500	337.00	393.08
1000	644.00	660.50
2000	1315.00	1438.57

at the 2000-gm. level. Some systematic factors such as size-weight illusions may have been involved at these levels. The general fit of the *AL*-distribution to the obtained distribution is so good, however, that the Harper-Stevens results may be explained on the assumption that their *S*s were actually making *AL*-judgments.

Guilford and Dingman recently studied psychological scales of weight constructed both by the method of fractionation and the method of constant sums.[8] Their curves vary with method, but in no case do their results substantiate the Harper-Stevens function. The general form of their curves based on fractionated data does agree, however, with the Harper-Stevens curve, as did the function reported earlier by Taback.[9] The use of some form of ratio-judgment method seems consistently to yield positively accelerated functions. Guilford and Cotzin reported a negatively accelerated curve in accord with the Weber-Fechner law in using a modification of the method of constant stimuli.[10]

Perhaps the experimental question which should be asked is not whether *the* psychological scale for weight is a positively or negatively accelerated function, but rather what it is about the conditions under which ratio-judgments are made that produces the one function, and what it is about the conditions that produces the

[7] Harry Helson, Adaptation-level as a frame of reference for prediction of psychophysical data, this Journal, 60, 1947, 1-29.

[8] J. P. Guilford and H. F. Dingman, A validation study of ratio-judgment methods, this Journal, 67, 1954, 395-410.

[9] Woodworth and Schlosberg, *op. cit.*, 243.

[10] J. P. Guilford and Milton Cotzin, Judgment of difficulty of simple tasks, this Journal, 54, 1941, 38-52.

other function when methods of successive intervals or constant stimuli are used.

The difference between the Guilford-Dingman and the Harper-Stevens functions lies in the slope of the curves. The Harper-Stevens curve is much steeper than any of those obtained by Guilford and Dingman. There are certain differences in stimulus-materials and procedure between the two situations that deserve further investigation. We would suggest (based on our results) that the use by Guilford and Dingman of a geometric comparison-series might account for the reduced positive acceleration of their scale in relation to that of Harper and Stevens, whose comparison-series were roughly linear. It would be of interest to determine if the *Ss* of Guilford and Dingman, too, were making bisectional judgments in terms of *AL*, but not enough information about their procedure was supplied to permit the computations.

Garner has shown recently that half-loudness judgments made with a method of constant stimuli are controlled by the context of stimuli presented to *S*.[11] Using three different ranges of variable intensities with the same standard, and using a separate group of *Ss* for each range, Garner found judgments which were very reliable but quite lacking in validity, since for a standard of 90 *db.*, the mean half-judgment by each group, respectively, was (rounded) 60, 70, and 80 *db.* The judgments, therefore, depended almost entirely upon the context, a finding which agrees wholly with our interpretation of the Harper-Stevens data.

Summary. An equal-interval scale of weights, constructed on the basis of data obtained by Harper and Stevens with the method of fractionation, was found to depart markedly from linearity, while scales constructed on the assumption that all *JNDs* are subjectively equal were found to approximate linearity very closely. A comparison was made between the Harper-Stevens obtained values of half-judgments and calculated adaptation-level values for each weight-level. Very high agreement was found, indicating that the method of fractionation may be a special case of the adaptation-level judgment. The method of fractionation is probably not a suitable method for constructing psychological scales of weight.

[11] W. R. Garner. Context effects and the validity of loudness scales, *J. Exper. Psychol.*, 48, 1954, 218-224.

LONG-RANGE CONSTRAINTS IN THE STATISTICAL STRUCTURE OF PRINTED ENGLISH

By N. G. BURTON and J. C. R. LICKLIDER, Massachusetts Institute of Technology

The sequences of letters that constitute printed English are among the most intricate and also among the most important patterns produced by human beings. As a characteristic of language behavior, the stochastic structure of letter sequences is therefore of special interest to psychologists.

Redundancy is an essential feature of the stochastic structure. The sequences are redundant in the sense that letters in one part limit the possibilities and influence the probabilities of letters in another part. If a literate subject is given a succession of 10 letters, the constraints of spelling, grammar, syntax, and idiom make it possible for him to predict with some accuracy what the eleventh will be.

The ordering of letters is so complex, actually, that the only statistical machine that can deal adequately with the problem is the human verbal mechanism. For this reason, although the study of redundancy in samples of printed English might in principle be restricted to counting letter sequences and making calculations on the basis of relative frequencies, practicable methods of investigation are as psychological as the problem.

Working on the assumption that an intelligent human being knows the probability structure of his language well enough to approximate the performance of an ideal predictor, Shannon developed a technique for making quantitative estimates of the redundancy of sequences of letters.[1] Showing *S* the first n-1 letters of an n-letter passage, he had *S* make 'educated guesses' until the correct nth letter was named. This procedure was repeated over and over with values of n running from 1 to 15 and also with $n = 100$. From the distribution of the number of guesses required in repeated trials by several *S*s, Shannon determined upper and lower bounds on the entropy or uncertainty or 'amount of information' in printed English.

The amount of constraint exerted by the context is best shown by the ***relative redundancy,*** which is one minus the ratio of the actual amount of information per letter to the amount (4.7 bits) that could be transmitted if the letters in the sequence were selected independently of one another. According to Shannon's data, a letter is, on the average, at least three-quarters determined by what has preceded it; less than one quarter of its information is new. The question is, is all that constraint upon the selection of a given letter imposed by letters near it in the sequence, or do distant influences have an appreciable effect. This is what is hard to tell from Shannon's data. The curves, plotted from Shannon's results to show relative redundancy as a function of the number (n) of letters in the test-sequence, rise noticeably between $n = 15$ and $n = 100$, but the course of increase between these

---

* Accepted for publication September 3, 1954. This research was supported in part under Air Force Contract AF 18(600)-322, monitored by the Operational Applications Laboratory, Air Force Cambridge Research Center, Air Research and Development Command.

[1] C. E. Shannon, Prediction and entropy in printed English, *Bell Syst. Tech. J.*, 30, 1951, 50-64.

two points cannot be determined, and the function therefore cannot be extrapolated. There is, consequently, no rigorous basis for deciding whether the estimate of redundancy based on $n = 100$ is a good figure for the *total* redundancy, or whether printed English might turn out to be, say, 95% redundant if very-long-range constraints were taken into account. Since long-range constraints include the influences of subject matter, style, level of presentation, and the dynamics of the situation reported or described, it is possible *a priori* that they might be quite strong. To estimate their strength—to obtain a more complete picture of the course of the curve relating estimated relative redundancy to sample length—is the object of the present study.

Procedure. We followed Shannon's experimental procedure closely except that our sources of test-material were 10 paper-backed novels instead of a single volume. (Shannon's only source appears to have been *Jefferson the Virginian* by Dumas Malone.) The 10 novels were of about the same level of reading difficulty (average $R\ E = 70.6$) and abstraction (average $R = 75.2$) as measured by Flesch's scales.[2] From each source we selected 10 passages of each of 10 lengths: $n\text{-}1 = 0, 1, 2, 4, 8, 16, 32, 64, 128$, and approximately 10,000 letters. In every case, the page, line, and position in the line of the test-letter were chosen with the aid of a table of random numbers.

Ten *Ss* were used, one for each novel. Each was either a graduate student or a psychological research technician. Considerable pains were taken with their selection, since we wished intelligent, highly verbal people who would approximate ideal predictors and who would be intrigued by the guessing game.

The *Ss* were told that they would be given passages of various lengths, removed from context, and that they should in each instance supply the next letter. They were instructed to count the space between two words as a letter. Selections of 128 letters or fewer were dictated to the *Ss*. They wrote these selections down and studied them before attempting to guess what came next. The 10,000-letter passages were presented as printed, with adjacent letters masked out. *Ss* worked individually, each at his or her own pace. In all cases they were required to continue guessing until they had named the next letter correctly. All guesses were recorded.

Results. Upper and lower bounds on relative redundancy, derived by Shannon's method from distributions of numbers of guesses in the 100 tests (10 samples × 10 *Ss*) for each value of n-1, are shown as solid lines and filled circles in Fig. 1. The open circles give the results of a retest to check the inversion in the data at $n\text{-}1 = 64$. Evidently the irregularity was a sampling fluctuation.

For passages more than four letters long, our estimates of redundancy are somewhat lower than those obtained by Shannon. Our results in the interval $n = 1$ to $n = 30$ are quite similar, however, to those of Frick and Sumby in tests with transcribed samples of radio messages between aircraft and control towers.[3] Certainly any differences in overall level of constraint can be ascribed to differences in type of material or differences among subjects.

[2] Rudolph Flesch, Measuring the level of abstraction, *J. Appl. Psychol.*, 34, 1950, 384-390.

[3] F. C. Frick and W. H. Sumby, Control tower language, *J. Acoust. Soc. Amer.*, 24, 1952, 595-596.

The curves of Fig. 1 level off at about n-1 = 32. The constraint imposed by 10,000 preceding letters is little greater than that imposed by 32. This does not necessarily mean that determining influences do not extend over long spans. It may be that knowledge of letters some distance removed will permit better-than-chance prediction. It appears, however, that remote preceding letters impose no marked constraints that are not imposed also by neighboring preceding letters. Written English does not become more and more redundant as longer and longer sequences are taken into account.

The confounding of *S*s with sources—the use of one novel with each *S*—was intentional but, as we now see it, unfortunate. In fact, the restriction of the source

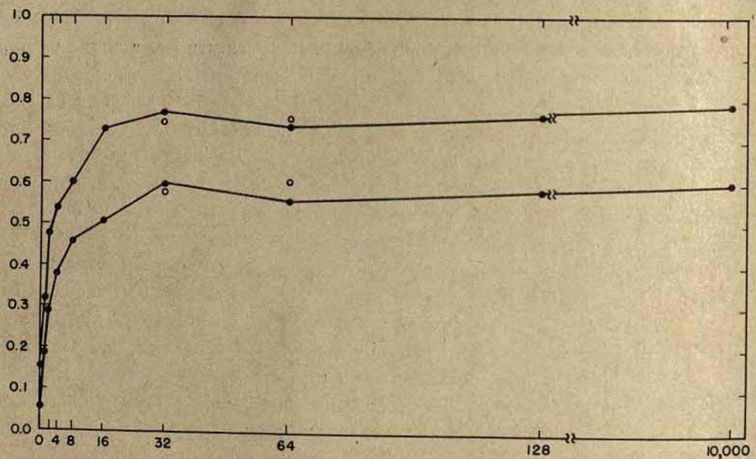


FIG. 1. ESTIMATED UPPER AND LOWER BOUNDS OF RELATIVE REDUNDANCY AS A FUNCTION OF THE AMOUNT OF PRECEDING TEXT KNOWN TO *S*s

material in any way that allows *S* to learn of the restriction is to some extent antagonistic to the aim of discovering how much long-range influences add to the constraint imposed by neighboring letters. It may be argued that much of what *S* gains from knowledge of the author's style or from the subject matter of the source he gains as soon as he sees or figures out what the source is. The effect of this argument is lessened, however, by the results of a postliminary experiment. Ten *S*s were able to predict the last letter of 10,001-letter sequences from known sources no better than the last letters of 33-letter sequences from comparable but unknown sources.

We may conclude, therefore, that constraining influences in printed English may preclude two-thirds or four-fifths of the freedom inherent in haphazardness, but that they do not approach complete determination much more closely than that, even when long-range forces are included in the measure. This conclusion may probably be extended with good approximation from letters to phonemes of the spoken language. It may be extended also, with at least fair approximation, to words and sentences, which are probably more natural units. The approximation in the latter

case is introduced by the fact that words and sentences are of unequal letter-lengths (or phoneme-lengths), but the non-uniformity probably does not bias the relative redundancy very much. In any event, the only approach to estimating the redundancy of word and sentence sequences appears to be through letters or phonemes. Shannon's guessing-game technique can in principle be applied to words and sentences, but it leads to trials that never terminate.

Summary. An experiment modeled after Shannon's was conducted to determine the extent to which estimates of the redundancy of English texts are dependent upon the number of preceding letters known to *S*. Data obtained indicate that, while the estimate of relative redundancy increases as knowledge of the foregoing text is extended from zero to approximately 32 letters, increasing the known number of letters beyond 32 does not result in any notable rise.

AUDITORY AND VISUAL FLICKER-FUSION AS MEASURES OF FATIGUE

By S. W. DAVIS, Johns Hopkins University

The fusion-frequency of visual flicker has shown considerable promise as a measure of fatigue.[1] Available evidence suggests that visual fusion is at least partially a central phenomenon, and that central fatigue lowers the fusion-frequency. In 1948, Miller and Taylor reported a series of experiments with interrupted (fluttering) white or gray noise—an auditory analogue of visual flicker—in which the rate of interruption was increased until the subject (*S*) reported fusion.[2] They found large individual differences, and (within normal intensity-ranges) a considerably higher fusion-frequency for audition *(AF)* than for vision *(VF)*. This work suggests the possibility that the state of the central nervous system may be reflected by auditory as well as by visual changes. For practical purposes, the measurement of *AF* would be more convenient, requiring simple apparatus, and presenting fewer problems such as control of the surround and the level of adaptation. The present study was designed to compare changes in auditory and visual fusion frequencies in the course of prolonged mental work.

Apparatus. The apparatus for measuring *VF* consisted of a DC light-source which was directed by means of a lens upon a reflector coated with magnesium oxide.[3] The reflector was placed at an angle of 45° immediately behind a small opening in a large white quarter-sphere which rested on a table at which *S* was comfortably seated. *S* viewed the reflector through this aperture from a distance of 18 in., at which the aperture subtended an angle of 2°. The light was caused to flicker by means of a rotating sector-disk driven by a DC motor, which interrupted the light-beam abruptly at the focal point of the lens. The speed at which the disk rotated was altered by means of variable resistors in the circuit. In all of the experiments reported here, the light was interrupted with an on-off ratio of 50%. When first presented, it was interrupted at such a rate that each *S* reported it as flickering. The rate was then increased until *S* could no longer perceive flicker.

The work-place was illuminated by frosted tungsten lamps, operated on DC, which were placed above and behind *S*. These lamps provided 53 ft.-c. of illumination on the work. The brightness of the surround was 43 ft.-L. The brightness of the flickering target was 40 ft.-L., an average value since it was measured while the light was being interrupted.

The noise-generator of the apparatus for measuring *AF* was a gas-tube which

---

* Accepted for publication November 29, 1954. This study, directed by Professor T. A. Ryan, is adapted from a doctoral dissertation presented to Cornell University.

[1] Ernst Simonson and Norbert Enzer, Measurement of fusion frequency of flicker as a test of fatigue of the central nervous system, *J. Indus. Hyg. & Toxicol.*, 23, 1941, 83-89; R. A. McFarland, A. H. Holway, and L. M. Hurvich, *Studies of Visual Fatigue*, 1942, 191-219; Joseph Brozek and Ancel Keys, Flicker fusion frequency as a test of fatigue, *J. Indus. Hyg. & Toxicol.*, 26, 1944, 169-174.

[2] G. A. Miller and W. G. Taylor, The perception of repeated bursts of noise, *J. Acous. Soc. Amer.*, 20, 1948, 171-182.

[3] The apparatus was constructed by Professors T. A. Ryan, M. E. Bitterman, and C. L. Cottrell with the aid of a grant from the Research Fund of the Illuminating Engineering Society.

produced a white noise. The output of the generator was fed into a 22-w., push-pull amplifier which had a frequency response of 60-10,000 c.p.s. flat within 3 db. The output of the amplifier was in turn fed into a switching amplifier, and from there into crystal earphones through an output-transformer. The earphones had a frequency response from 100-9000 c.p.s. The intensity of the noise at the earphones was controlled at the amplifier, and was fixed at a comfortable level which was the same for all *Ss*.

The white noise was interrupted by applying negative voltage to the screen-grids of the tubes in the switching amplifier at fixed intervals. The application of this voltage was controlled by the output of a multivibrator, with a continuous range of 50-400 c.p.s. The on-off ratio was 90%.[4]

When first presented, the noise was interrupted at 50 c.p.s., which was below the fusion-frequency for all *Ss* and the rate of interruption was then increased slowly until *S* reported the disappearance of flutter. The physical surround was identical with that used for measuring *VF*.

Procedure. The procedure was, in general, the same for each of the experimental series. First, *S* was instructed on how to report the perception of fusion to *E*, and a series of three practice-measurements was made in each modality. Secondly, formal determinations of *VF* and *AF* were made, with each determination based on 10 measurements and with the order (visual-auditory or auditory-visual) being alternated for each succeeding *S*. Thirdly, *S* was presented with a series of multiplication-problems, each consisting of a pair of two-digit numbers. The numbers were selected from a table of random numbers, omitting zeros and double-digits. They were arranged in columns of 35 on legal-size paper with a space left for the answer to the right. They were divided into groups of 5 problems by lines drawn horizontally on the paper. *S*'s task was to multiply these pairs of digits mentally as rapidly and as accurately as possible.[5] Finally, *VF* and *AF* were determined again in the same manner and in the same sequence as before the work. *S*'s instructions were as follows.

> You will notice on the paper a series of pairs of two-digit numbers. Your task is to multiply them in your head. To insure that you do them in your head I am asking that you write nothing on the paper but your answers. You may draw imaginary figures in the air, talk out loud, or use any short-cut or trick you want. There are 140 (or 70 depending on the experimental series) problems on the paper; however, that selection is arbitrary. Your task is to work as rapidly and as accurately as you can for 2 hr. [1 hr. in Series 2]. If you should complete the problems on the sheet in less than this time, I have more problems for you to do. If you should do less than 140 in the two hours, that is perfectly all right. Are there any questions?

S was then told to begin work, and while he worked, the time necessary to complete each group of five problems was recorded.

RESULTS

Series 1. In this series, the effect of 2 hr. of mental arithmetic on *AF* and *VF* was measured. Twenty *Ss*, all University students, were studied, and the task-sheet contained 140 multiplication-problems. The results are summarized in Table I.

---

[4] The apparatus was designed by Mr. Wilson Greatbatch.

[5] This task is similar to one used by Tsuru Arai (*Mental Fatigue*, Teachers College, Columbia University Contributions to Education, No. 54, 1912, 1-92).

Significant decreases both in *AF* and *VF* appeared, although percentage-change in *AF* was significantly greater than that in *VF* at the 5% level of confidence ($t = 2.54$ with 19 *df*). The correlation between change in *AF* and change in *VF* was found to be 0.40, a value just short of the 5-% level of significance with 19 *df*. It is interesting to note also that there was no relation between change in fusion-frequency and the number of problems accomplished during the working period. The correlation was 0.09 for *AF* and 0.10 for *VF*, with both coefficients falling far short of statistical significance.

TABLE I

MEAN *AF* AND *VF* BEFORE AND AFTER TWO HOURS OF MENTAL MULTIPLICATION

	Before	After	Diff.	t	P
AF	177.5	159.7	−17.8	3.50	0.01
VF	36.0	35.2	− 0.8	2.25	0.05

Series 2. In this experiment, conducted under the same conditions as the first, the working time was reduced to 1 hr. One group (A) of 20 *Ss* was presented with the 140 problems used in the previous experiment, while a second group (B) of 20 *Ss* was presented with only 70 problems. Although *E* did not instruct *S* to solve any given number of problems or press him in any way, it was felt that *S* would attempt an accomplishment in relation to the number of problems presented to him. Thus, by presenting 140 problems and cutting the working time in half, the pressure to exert effort would be as great or greater than it was in Series 1, while the actual work done would be less. For Group B, given 70 problems rather than 140, both the task and the time were reduced. The results for Series 2 are shown in Table II.

TABLE II

MEAN *AF* AND *VF* BEFORE AND AFTER ONE HOUR OF MENTAL MULTIPLICATION AS A FUNCTION OF THE NUMBER OF PROBLEMS PRESENTED

	Number of problems	Before	After	Diff.	t	P
AF	70	201.8	192.7	−9.1	3.14	0.01
	140	207.2	200.5	−6.7	3.68	0.01
VF	70	37.3	37.9	0.6	2.01	0.10
	140	37.4	37.5	0.1	0.85	—

The number of problems presented seemed to make no difference. In both groups *AF* declined significantly in the course of an hour of work while *VF* did not. Apparently *VF* provides a less sensitive measure of fatigue than does *AF*. The decline in *AF* was smaller after 1 hr. of work than after 2 hr., the difference approaching significance at the 6-% level when data for the 40 *Ss* of the second experiment were compared with data for the 20 *Ss* of the first ($t = 1.90$). For the 60 *Ss* of the two series combined, there was no significant correlation between changes in *AF* and *VF*. The coefficient, corrected for co-variance in the three different experimental conditions, was 0.21 ($p > 0.05$). Nor was there a significant relation between pre-work *AF* and *VF* scores for all *Ss*—the computed correlation was 0.22 ($p > 0.05$). It would seem then, that although both auditory and visual flicker appear to reflect changes due to work, they do not vary together in a given individual.

The results suggest that *AF* is a more stable indicator of fatigue than is *VF*. It

should be pointed out, however, that intensities and on-off ratios other than those used in this experiment might very well provide different results. A further exploration of the effect of these factors is needed.

Summary. Decreases in the fusion-frequencies of visual flicker and auditory flutter were found in the course of 2 hr. of mental arithmetic, the latter change being more marked than the former. During 1 hr. of mental arithmetic, only the auditory measure declined. The results suggest that auditory flutter-fusion may provide a sensitive and convenient index of fatigue.

WHOLE AND PART METHODS IN LEARNING A PERCEPTUAL MOTOR SKILL

By F. J. McGuigan, Hollins College, and Eugene F. MacCaslin, George Washington University

There has been little study of the relative efficiency of whole and part methods in the learning of complex, perceptual motor skills. The only directly relevant experiment which has produced statistically reliable results is one by Barton, who found that typewriting is better learned by the whole than by the part method.[1] The skill investigated in the present study, rifle marksmanship, is similar to typewriting in that both are complex, perceptual motor skills; neither can be adequately learned in a day, and both may be more or less arbitrarily analyzed into several subtasks, each of which bears a meaningful relationship only to its whole. The primary purpose of the study was to determine whether a whole method or a part method is superior for learning rifle marksmanship.

A secondary purpose of the study was to determine whether the relative efficiency of whole and part methods for the learning of perceptual motor skills is a function of intelligence. Previous research has shown that "the higher the level of mental development, the more likely is the whole method to be superior."[2] The study also was concerned with the effect of massing practice on an important subtask (firing, the goal-act of skill in marksmanship), as in the part method, versus distributing it, as in the whole method. Practice by the whole method on any particular subtask is necessarily distributed from rehearsal to rehearsal, while practice by the part method on that subtask tends to be massed (regardless of whether *training periods* are massed or distributed).

Method

Subjects. The first experiment was conducted with 148 trainees of a Fort Knox, Kentucky, infantry company; the second experiment (a repetition of the first) used 200 trainees of a Fort Jackson, South Carolina, infantry company. The *S*s in each experiment were beginning Army basic training.

Training procedure. Several subtasks comprise the process of learning to fire an Army rifle. The trainee must learn (a) to assume certain well-defined postures for firing in the standing, kneeling, sitting, and prone positions; (b) either to wrap ('hasty sling') or loop ('loop sling') the rifle sling about his arm as the position demands, to give him greater stability; (c) to move the sight-adjustment to alter the strike of the bullet; (d) to load and unload his weapon; (e) to align his sights on the target to obtain the proper 'sight picture'; (f) to control his breathing—to take a breath and hold it until he has fired; and (g) to squeeze the trigger gradually (to avoid stimuli to which an anticipatory startle or flinching response might be

* Accepted for publication December 3, 1954. The research reported here was conducted by the authors while they were employed by the George Washington University, Human Resources Office, operating under a contract to the Department of the Army. Opinions and conclusions are those of the writers and do not necessarily represent views of the University or the Department of the Army.

[1] J. W. Barton, Comprehensive units in learning typewriting, *Psychol. Monogr.*, 35, 1926 (No. 164), 1-47.

[2] J. A. McGeoch and A. L. Irion, *The Psychology of Human Learning*, 1952, 501.

anchored). After he has learned to do these things in slow-fire instruction, he must learn to adapt the same techniques to sustained (rapid) fire.

In both experiments, *Ss* were trained in marksmanship according to a 28-hr. schedule (training periods 1 to 4 hr. in length) over a period of two weeks. About 20 of these hours were devoted to slow fire instruction. *Ss* were matched on the basis of their scores on an ataxiametric test of rifle steadiness,[3] and assigned to the following groups.

Group I (Part Method). This group was trained by a repetitive part method, receiving instruction on the first subtask of the firing act, then on the second simultaneously with practice on the first, then on subtasks (a), (b), and (c) together, and so on until all were put together into the act of firing. Firing was massed during the last 4 hr. of training.

Group II (Whole Method). In the whole method of training used in this study, *S* first watched a half-hour demonstration of the entire sequence of movements entering into the act of firing. Subsequently, *in each training period* of the 28-hr. schedule, he received both instruction and practice on all of the subtasks, including firing.

Group III (Incomplete-Whole Method). One aspect of the act of firing a rifle requires *S* to learn a set of responses to its loud report and recoil. Firing also furnishes knowledge of results (target-information) and therefore affects firing performance. To determine the effect of massing practice on this important subtask (goal-act) at the end of training (as in the part method) as compared with distributing practice on it throughout training (as in the whole method) an incomplete whole method was introduced into the design of the study. This method was used for Group III. In the first 24 hr. of the 28-hr. schedule, the training of Groups II and III was identical except that whenever Group II fired, Group III simulated firing. In the last 4 hr. the 'live' firing of Group III, like that of the part group, was massed.

Group IV (No Training). This group received only essential instruction, which included training in safety precautions and range procedure. Using this group made possible a more adequate evaluation of the results of training.

The four groups used in each experiment were not significantly different in previous experience with weapons (as measured by a questionary), in handedness, in frequency of extreme defects in visual acuity, in intelligence (as measured by scores on Army Classification Battery Aptitude Area I), or in number of years of education. Groups I, II, and III spent the same number of hours (28) in training, and each group fired the same number of rounds. Group IV was concurrently given subject matter not related to weapons-skill. To facilitate administratation, each group comprised a platoon. The platoons were housed separately to reduce communication between them.

Testing procedure. The data were the scores obtained during four days of firing on an Army range, each *S* firing a total of 100 rounds in slow fire and 72 rounds in sustained fire at distances of 100, 200, 300, and 500 yd.[4] These scores were

---

[3] Previous studies show a high relationship between steadiness and marksmanship. The average correlation between steadiness and scores in slow fire obtained in this study was only −0.24 (the ataxiameter measured unsteadiness); the correlation between steadiness and sustained fire was insignificant. See F. J. McGuigan and E. F. MacCaslin, The relationship between rifle steadiness and rifle marksmanship and the effect of rifle training on rifle steadiness, *J. Appl. Psychol.*, 35, 1955, 156-160.

[4] Practical considerations made it necessary to omit 500-yd. slow firing in the

recorded in terms of target hits: 5 points for a bullseye and 4 or 3 points for hits in concentric rings. The points for each *S* were summed and the data were analyzed to provide *two criteria,* slow fire and sustained fire.[5]

RESULTS

Reliability of criteria. The estimated reliability of the slow fire was 0.88 for the first experiment and 0.84 for the second. The reliability of the sustained fire was 0.83 for the first experiment and 0.81 for the second.[6]

Whole versus part. The general pattern of results is shown in Table I. Precision of measurement was increased by using an analysis of covariance to equate the groups for intelligence. Both the part and whole groups were significantly superior to the no-training group beyond the 1-% level, on both criteria. An adjusted *F*-test showed that neither the part nor the whole method is superior for sustained fire training (adjusted $F = 0.002$ and 1.22, respectively, for the two experiments). When, however, the levels of significance reached in the two experiments are compounded,[7] the whole method is superior to the part method beyond the 5-% level of significance (adjusted $F = 5.33$ and 2.14, respectively) for slow fire training.

Effect of intelligence. The groups were divided (at the theoretical mean of 100) into two subgroups, above-average (100 and above) and below-average (99 and below) in terms of scores on Army Classification Battery Aptitude Area I. The interaction between methods (part and whole) and intelligence did not reach statistical significance either for slow or for sustained fire. When the probabilities for sustained fire were compounded, however, significance was approached (9-% level, $F = 1.59$ and 3.09 for the two experiments). The direction of the differences (see Table I) suggests that for *slow* fire the whole method is superior to the part method for *Ss* of both below-average and above-average intelligence; whereas for *sustained* fire, superiority of whole method training may be evident only with the above-average *S*.

Massing versus distributing of a goal-act. It is possible to view a comparison

second experiment. Thus *Ss* in the second experiment fired 74, rather than 100, rounds in slow fire. The number of rounds of sustained fire (72) was the same for both experiments. An incidental finding of this study was that the differences between methods increased as a function of the distance of the target. The fact, reported below, that the probability-level for the slow fire difference between whole and part methods dropped from the first to the second experiment may thus be partially due to the omission of 500-yd. firing in the second experiment.

[5] Slow and sustained fire have some qualitative differences. In slow fire, *S* has an unlimited time between shots and is given target information after each shot. In a sustained fire, he has 50 sec. in which to fire 9 rounds and change a clip of ammunition. He receives no information about his accuracy until the end of the exercise. Another distinction lies in the fact that in sustained fire the rather pronounced recoil of the weapon requires that the trainee become adept in rapidly readjusting the alignment of his weapon. With experienced firers, this repeated movement becomes quite rhythmic.

[6] These coefficients were obtained with the use of the Spearman-Brown formula. The same set of firing exercises, with allowances in the analysis for minor differences, was administered three times during the four days of firing. The correlations among these trials for each criterion were found to be not significantly different from each other within each experiment. Fisher's *z*-method was then used to obtain an average coefficient in each case. The average coefficient was used in the Spearman-Brown formula, with $n = 3$, to yield the coefficients reported here.

[7] E. F. Lindquist, *Statistical Analysis in Psychological Research,* 1940, 46-47.

between the incomplete-whole method and the part method as a comparison between the whole method and the part method, the kind of *firing* (massed) being held constant. A comparison between the incomplete whole method and the whole method is, essentially, a comparison between distributed and massed firing, the kind of *method* (whole) being held constant. For sustained fire, no significant difference

TABLE I

PERFORMANCE AS A FUNCTION OF TRAINING METHOD FOR Ss OF HIGH AND LOW INTELLIGENCE AND FOR ALL Ss COMBINED

Method		Scores in slow fire			Scores in sustained fire		
		Ss of high intel.	Ss of low intel.	Ss combined*	Ss of high intel.	Ss of low intel.	Ss combined*
Part	*M*	246.3	215.2	222.6	155.9	132.7	138.2
	SD	70.5	76.4	75.4	58.6	64.1	62.9
	N	21	65	87	21	65	87
Whole	*M*	299.8	217.4	248.8	189.8	125.0	148.9
	SD	70.9	65.3	77.8	53.4	50.2	60.6
	N	33	53	87	33	53	87
Incomplete whole	*M*	265.9	212.1	228.7	176.5	123.2	138.3
	SD	78.8	65.6	75.0	43.1	52.0	56.0
	N	21	64	87	21	64	87
Control	*M*	247.0	163.7	188.8	149.0	94.4	110.5
	SD	54.2	69.6	76.4	49.4	55.4	58.9
	N	23	62	87	23	62	87

* *N*s for combined *S*s are larger than those for sums of high and low *S*s because a few intelligence scores were not obtained. Tests between methods were based on *N*s of 87

was found between massing or distribution of firing. A tentative conclusion about slow fire may, however, be drawn. When the probabilities for the two experiments are compounded, the incomplete-whole method and the part method do not differ significantly (50-% level, $t = 1.19$ and 0.27). The incomplete-whole method and the whole method, compared in the same way, differ at a level of significance beyond 8-% ($t = 1.52$ and 1.67). Although these data are not conclusive, they suggest: (1) that the incomplete-whole method and the part method (both of which use massed firing) do not differ significantly from each other; (2) that the whole method (which uses distributed firing) may be superior to both of the other methods. It is therefore possible that much of the superiority of the whole method over the part method in slow fire is due to the fact that the whole method involves distributed practice of the goal-act.

CONCLUSIONS

With reference to the three purposes of the study, the following conclusions have been reached. (1) The whole method of learning a complex, perceptual motor skill, such as rifle marksmanship, is superior to the part method in slow-fire for *S*s of all levels of intelligence. (2) The superiority of the whole method over the part method in sustained-fire approaches significance for *S*s of above-average intelligence. (3) Although the finding fails to reach significance, the data suggest that much of the superiority of the whole method over the part method for slow-fire training may be due to the fact that the whole method involves distributed practice in firing, the goal-act of rifle marksmanship.

ASSOCIATIVE FREQUENCY* AND ASSOCIATIVE PREPOTENCY AS MEASURES OF RESPONSE TO NONSENSE SYLLABLES

By GEORGE MANDLER, Harvard University

The extensive use of nonsense syllables in the study of verbal behavior has repeatedly led to a concern with the measurement of 'meaning' of such stimuli. Indices of the meaningfulness of nonsense syllables have been developed by Cason,[1] Glaze,[2] Hull,[3] Krueger,[4] and Witmer.[5] More recently Noble has briefly reviewed this work and derived a new measure, '*m*.'[6] The present study reports the development of two indices for nonsense syllables related to the work of Glaze, Krueger, and Noble.

Glaze and Krueger were primarily concerned with the determination of the *associative value* of nonsense syllables. This value was defined as the tendency of *Ss* to ascribe meaning to nonsense syllables, *i.e.* it is a measure of the extent to which nonsense syllables elicit 'ideas' or 'meanings' associated with them. In subsequent work it has been shown that this index is a useful predictor in situations involving verbal learning.[7]

Noble has attempted to define the concept of meaning in terms of Hullian theory. He suggests that 'meaning' be "coördinated with Hull's theoretical construct *H* by postulating that meanings increase as a simple linear function of the number of *S*[1]*-multiple R* connections acquired in a particular organism's history." He further defined the index of meaning, *m*, as the "mean frequency of continued written associations made by *Ss* within a 60-sec. time-interval."[8] The usefulness of such an index has been borne out by Noble's subsequent work.[9]

In a recent discussion of response-factors in human learning, we have suggested a line of reasoning somewhat similar to Noble's.[10] We have assumed that all stimuli, and responses functioning as stimuli, with which *S* has had previous experience, or which lie on a similarity continuum with such previously experienced stimuli, will evoke one or more differentiating responses. These responses will differ from each other in respect to their probability of evocation. We will describe here two indices of response for nonsense syllables which are related to that theoretical framework; namely, (a) the mean frequency of continued association evoked during a fixed

* Accepted for publication October 13, 1954. This research was supported by the Laboratory of Social Relations, Harvard University. The valuable assistance of Janellen Huttenlocher, Shirley Shapiro, and Peggy Charren is gratefully acknowledged.

[1] Hulsey Cason, Specific serial learning: a study of backward association, *J. Exper. Psychol.*, 9, 1926, 195-227.

[2] J. A. Glaze, The association value of non-sense syllables, *J. Genet. Psychol.*, 35, 1928, 255-267.

[3] C. L. Hull, The meaningfulness of 320 selected nonsense syllables, this JOURNAL, 45, 1933, 730-734.

[4] W. C. F. Krueger, The relative difficulty of nonsense syllables, *J. Exper. Psychol.*, 17, 1934, 145-153.

[5] L. R. Witmer, The association value of three-place consonant syllables, *J. Genet. Psychol.*, 47, 1935, 337-359.

[6] C. E. Noble, An analysis of meaning, *Psychol. Rev.*, 59, 1952, 421-430.

[7] *Cf.* J. A. McGeoch, The influence of associative value upon the difficulty of nonsense-syllable lists. *J. Genet. Psychol.*, 37, 1930, 421-426.

[8] Noble, *op. cit.*, 429.

[9] Noble, The role of stimulus meaning (*m*) in serial verbal learning, *J. Exper. Psychol.*, 43, 1952, 437-446.

[10] G. Mandler, Response factors in human learning. *Psychol. Rev.*, 61, 1954, 235-244.

time-interval (to be referred to as *associative frequency* or f), and (b) the tendency of a stimulus to evoke the same association from different *Ss* (to be referred to as *associative prepotency or* p).

The former index is practically identical with Noble's m, while the latter describes presumably different properties of the stimulus. The term 'associative' is used only in a descriptive sense and does not involve any theoretical presuppositions or implications. Nor are we concerned with designating either measure as representative of the 'meaning' of the stimulus. It is suggested, however, that each of these definitions includes some of the operations usually associated with such 'meaning.'

Procedure. A sample of 100 nonsense syllables was selected from Glaze's lists.[11] To secure a wide range of syllables, they were chosen from lists having associative values of 0, 20, 40, 60, 80, and 100%, respectively. Sixteen syllables were selected at random from each of these lists, with an additional syllable from each of the 20, 40, 60, and 100% lists to make a total of 100 syllables. Only syllables that were likely to evoke strong emotional or blocking responses (three in all) were excluded from this selection.

The stimulus-items were printed in block letters approximately ¼ in. high in the center of 4¼ by 5½-in. white sheets. Every *S* received a test-booklet with the 100 stimulus-sheets so arranged in random order that no *S* responded to the syllables in the same order.

Every *S* responded to all stimulus-items in one session. Preliminary tests indicated that a 30-sec. exposure-period was sufficient to elicit most readily available associations. The *Ss* were 34 paid volunteer undergraduate students divided for testing purposes into three groups.

Instructions. The *Ss* were instructed that the experiment was designed to find out "what kind of associations people give to a series of syllables." They were told to write on each page of the test-booklet, *i.e.* around the syllable, all the words that occurred to them, or which they thought of, as they looked at, thought of, or pronounced the syllable. They were specifically warned not to give chain-associations. *Ss* were asked to write down as many words as they could think of during the time allowed. *E* then indicated to the *Ss* when to start work on the first syllable, stopped them after 30 sec., instructed the *Ss* to turn to the next page in the booklet, and so forth.

Results. Two indices (f and p) were obtained for each syllable. The index of associative frequency f was determined as the mean number of associations elicited by each syllable. The index of associative prepotency p was determined by establishing a frequency count of all specific associations elicited by each syllable. An arc sin transformation of the percentage of *Ss* who gave the most frequent association represents the value of p. Table I shows an alphabetic list of the 100 syllables with the values of f, standard deviation of f, value of p, and the most frequent association elicited by each syllable (the prepotent response). The 100 f-values vary from 2.9 to 5.3 (mean $= 4.1$, $SD = 0.48$), the p-values range from 24 to 76 (mean $= 46$, $SD = 9.1$).

The question now arises of the relationship between these two indices, and with the associative values established previously. Since the 100 nonsense syllables were selected from six of Glaze's lists (with associative values of 0, 20, 40, 60, 80 and 100%), an analysis of variance of the between-list variance of the two indices was performed. F-values for both indices are significant at better than the 0.001-level of significance. Both f- and p-values also increase as a function of Glaze's values, and the linear component in the analysis of variance is significant at better than the 0.001-level for both measures.[12] It will be recalled that both Glaze's and Krueger's associative values are based on the percentage of *Ss* who give a meaningful response to the nonsense syllables. Since Glaze's value are rectangularly distrib-

---

[11] Glaze, *op. cit.*, 262-265.
[12] W. G. Cochran and G. M. Cox, *Experimental Designs*, 1950, 58.

uted, a correlational analysis is not possible. Furthermore, while Glaze used only 15 *Ss*, Krueger's presumably comparable values are based on the responses of 200 *Ss*. An arc sin transformation of Krueger's associative values results in an approximately normal distribution of our 100 syllables. The product-moment correlations

TABLE I

RESPONSE-MEASURES OF 100 SELECTED NONSENSE SYLLABLES

Nonsense syllable	f	*SD* of f	p	Prepotent association*	Nonsense syllable	f	*SD* of f	p	Prepotent association*
BAL	4.9	1.53	62	BALL	POG	4.3	1.70	56	POGO
BAW	4.5	2.27	48	BAWL	PUQ	3.6	1.89	45	PUCK
BIP	4.3	1.94	29	BLIMP, BIB	QEZ	3.9	1.78	45	QUIZ
BIZ	4.2	1.76	72	BUSINESS	QID	4.2	1.48	44	QUID
BOM	4.2	2.05	59	BOMB	QIJ	3.4	1.71	43	QUIZ
BUV	4.0	1.88	39	ABOVE	QIP	4.0	1.78	55	QUIP
CEY	4.2	1.50	53	CEYLON	QUG	3.8	2.00	25	MUG
CID	4.5	1.79	34	KID	RAV	4.5	1.89	63	RAVE
CIN	4.9	1.76	70	SIN	RIL	4.9	1.79	36	RILE
CIX	3.8	1.90	50	SIX	RUV	4.0	2.07	31	ROUGH
CUW	4.0	1.67	34	COW	SAJ	3.9	1.83	32	SAGE
DAX	3.8	1.94	32	TAX	SAQ	3.6	1.86	57	SACK
DIF	3.8	1.68	40	DIFFERENCE	SAR	4.2	1.46	35	SARGENT
DIT	4.3	1.73	57	DITTO	SAX	4.7	2.32	64	SAXOPHONE
DIW	4.2	1.94	48	DEW	SIB	3.8	1.63	34	SIBLING
DUJ	3.7	1.73	31	DUZ	SIJ	3.6	1.90	26	SICK, SEDGE
FEL	4.5	1.89	48	FELL	SIK	4.4	2.14	55	SICK
FEY	4.5	1.89	31	FAY	SUB	4.7	1.42	59	SUBMARINE
FIJ	4.2	1.42	51	FIDGET	TEF	3.6	1.71	43	TOUGH
GAN	3.9	1.92	43	GAIN	TEX	4.8	1.65	59	TEXAS
GEW	3.8	1.71	42	JEW	TUS	4.2	1.30	48	TUSSLE
GIV	4.0	1.97	76	GIVE	VAJ	3.4	1.58	29	VAT
GIY	3.6	1.83	53	GUY	VEC	3.7	1.32	48	VECTOR
GOS	4.5	1.98	45	GOOSE	VEW	4.1	1.77	70	VIEW
GUV	3.6	1.31	42	GOVERNMENT	VIN	5.0	1.32	57	WINE
HAV	4.1	1.29	63	HAVE	VOJ	3.1	1.70	36	VOTE
HIY	4.0	1.98	48	HI (HY)	VOX	4.0	1.59	54	VOICE
HUV	3.9	1.72	39	HAVE	WAT	4.8	1.60	54	WATER, WHAT
JAT	4.4	2.16	39	JAP	WEF	4.1	2.01	29	FEW
JEK	4.1	1.61	39	CHECK, JERK	WIJ	3.7	1.59	31	WEDGE
JOH	4.5	1.71	56	JOHN	WUF	3.9	2.04	45	DOG
JUS	4.4	1.68	57	JUSTICE	WUH	4.3	2.07	41	WHAT
KEF	3.5	1.28	24	KEPT, KEEP	XAP	4.4	2.30	38	SAP
KOV	3.5	1.74	34	COVER	XEH	3.6	1.95	33	HEX
LAJ	3.7	1.70	36	LATCH	XIS	3.7	1.70	29	EXIST, KISS
LAT	5.3	1.66	43	LATE	XIV	3.1	1.58	33	14
LIM	4.7	1.60	54	LIME	XOK	2.9	1.79	29	SOCK
LIR	4.7	1.81	46	LYRE	XUR	3.4	1.61	29	SURE
LIX	4.2	1.77	47	LICKS	YAP	4.6	2.19	43	DOG
LUY	3.6	2.18	42	LIE	YAS	4.0	1.97	72	YES
MAB	4.1	1.56	36	MABEL	YEG	4.2	1.93	53	EGG
MOH	4.5	1.85	43	MORE	YEM	4.2	1.73	29	YUM, YAM
MON	5.1	1.78	57	MONDAY	YOV	3.8	1.93	36	YOU
NAD	4.1	1.81	37	MAD	YUN	4.2	2.03	52	YOUNG
NAZ	4.0	1.52	34	NAZI	ZAN	4.0	1.49	34	ZANY
NIS	4.0	2.13	40	NICE	ZEG	3.5	2.09	31	KEG
NIY	3.6	1.71	40	KNEE	ZIC	3.5	1.46	44	SICK
NUH	4.2	1.58	46	NO	ZOP	4.2	2.01	33	SOP
PAS	5.3	1.73	53	PASS	ZUK	3.7	1.91	33	SUCK
PAX	4.5	1.94	61	PEACE	ZUV	3.1	1.67	29	LOVE

* Where two different associations were given by the same number of *Ss*, both of these prepotent associations are listed.

between these transformed Krueger values and f and p are 0.651 and 0.724 respectively. The correlation between f- and p-values is 0.483. The correlations are highly significant. The relatively high correlations found for both of our indices with Krueger's associative values, coupled with the lower intercorrelation between f and p, suggest that the previously used associative values can be partly accounted for by

these two new measures. In other words, meaningfulness as defined by Glaze and Krueger does not seem to be a unitary dimension of nonsense syllables.

A comparison of our *f*-values with Noble's data for 96 dissyllables shows that our values all fall within the third highest quartile of his distribution. Thus, our *Ss* produced as many associations to nonsense syllables in 30 sec. as Noble's *Ss* produced in response to actual words in 60 sec. It is suggested that this difference is primarily associated with sampling differences. Our *Ss* were college students while Noble used military personnel.

The utility of the two measures developed here depends on further research. As we have suggested above, these indices were developed from some theoretical considerations concerning human learning. We would suggest that those syllables which evoke a large number of associations are more easily associated with each other than those syllables which evoke fewer associations, and furthermore that individual *Ss* who give many associations to nonsense syllables will show faster learning of new nonsense syllables than those *Ss* who give few associations. Experimental investigation of these propositions is now in progress.

APPARATUS

A DEVICE FOR OBSERVING ANIMALS IN DARKNESS

By ROBERT R. COX and LAWRENCE KRUGER,
Institute of Living, Hartford, Connecticut

The training of animals to perform somesthetic discriminations in the absence of visual cues has proven to be a long process, even in experiments with primates. The difficulty stems in large part from the need to conceal the discriminanda from view of the animal, thereby forcing the animal to place his arm in an awkward position to reach the objects. One way of obviating this difficulty is to place the discriminanda in a more comfortable position and to exclude visible light. The situation is analogous to the difficulties that a human *S* encounters in performing manipulations behind his back which can be accomplished with ease in a more comfortable position, even in the dark. The purpose of this apparatus, therefore, is to arrange conditions for observing animals performing various manipulations in the dark that they may be more easily trained for behavioral studies in somesthesis. Our experience with this apparatus to date indicates a considerable reduction in the number of trials necessary for a monkey to learn a somesthetic discrimination of form.

Two methods of observing animals in darkness were studied. The first method considered was a tube, popularly known as the "snooperscope," that converted the infra-red image.[1] Our experience with the British (E.M.I. Ltd.) type of tube was disappointing because of the large amounts of infra-red radiation needed, poor picture definition, and instability of the tube. The problem of flooding the animal with too much heat is of extreme importance in view of the deterioration of performance in a hot area.[2] Alternatively, a simple electro-mechanical scanning device was constructed which utilizes the principle of the Nipkow disk. This apparatus has proven to be stable and has required minimal maintenance.

Technical description. The device consists of a 'flying spot' of infra-red

---

* The apparatus described here was developed under contract with the U. S. Army, DA-49-001-MD-401.

[1] G. A. Morton and L. E. Florey. Infra-red image tube, *Electronics*, 19, 1946, 112-114.

[2] K. H. Pribram. Some physical and pharmacological factors affecting delayed response performance of baboons following frontal lobotomy. *J. Neurophysiol.*, 13, 1950, 373-382.

radiation which scans the field in which the discriminanda are placed across 32 lines at 20 times per sec., thereby preventing continuous flooding of the entire field with heat. The light source consists of a 6-v., 100-cp., DC operated bulb mounted in a modified slide projector. In place of the slide there is a 16-in. metal disk perforated with 32 small holes of equal diameter arranged in a spiral by increments of the hole diameter spaced 11.25° apart,

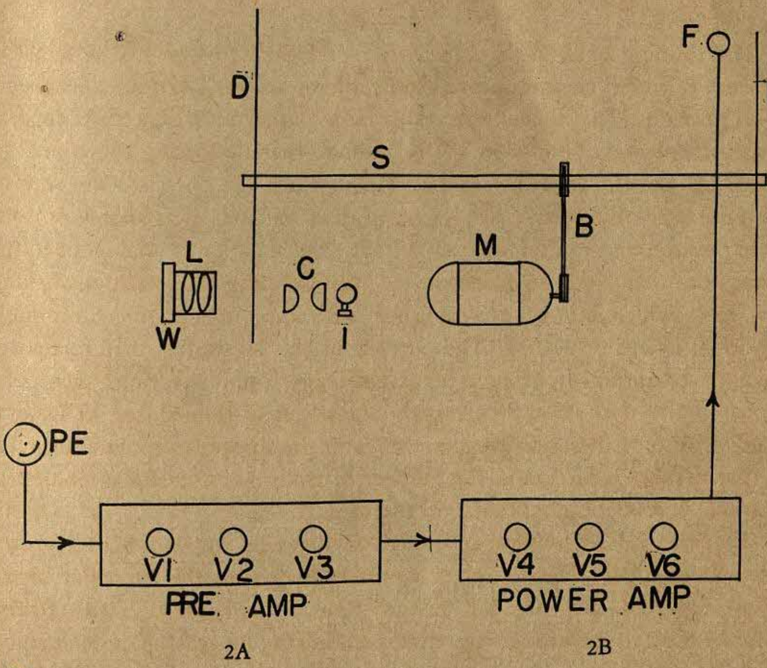
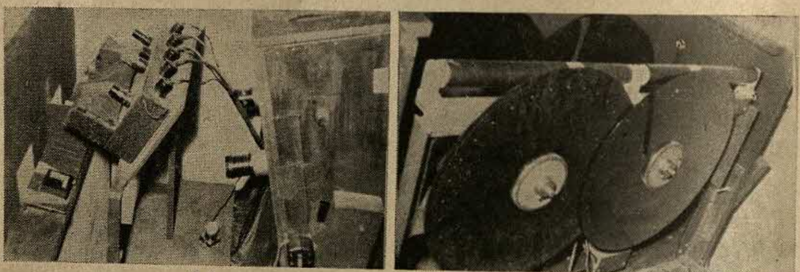


FIG. 1. BLOCK DIAGRAM OF ELECTRO-MECHANICAL INFRA-RED SCANNING DEVICE
B = Pulley belt; C = condensor lens; D = scanning disk; D′ = viewing disk; F = 14-w. fluorescent lamp; I = 6-v. incandescent lamp; L = projector lens; M = driving motor; PE = photoelectric cell; S = shaft; V1, = 6SJ7 vacuum tube; V2 and V3 = 6SL7 vacuum tube; V4 = 6J5 vacuum tube; V5 = 6V6 vacuum tube; V6 = 6L6 vacuum tube; W = Wratten No. 87 filter.

in close proximity to the edge of the disk. The disk is mounted on the front end of a shaft which rotates at 20 r.p.s. driven by a non-synchronous motor. In front of the disk is the projector lens which is covered with a Wratten No. 87, infra-red transmission filter, shown on the right of Fig. 2A.

The infra-red radiation which is scanning the field is in turn reflected onto an infra-red sensitive photocell (RCA Type 918) which converts the light energy into an electric current proportional to the incident light on

the cell. This small fluctuating current is amplified by a frequency compensated 3-stage battery-supplied preamplifier, the output of which, in turn, is introduced into a 3-stage power amplifier. The position of the photocells and preamplifiers with relation to the stimulus-boxes can be seen in Fig. 2A. The cathode follower output of the power amplifier operates a conventional 14-w. fluorescent lamp whose light output varies in proportion



2A 2B

FIG. 2. TWO PICTURES OF THE SCANNING DEVICE

In Fig. 2A the box at the right contains disks and lamps, the projection-lenses are at the center, the photo-electrical cells with pre-amplifiers are supported over the stimulus-trays. In Fig. 2B, two viewing disks are shown in the foreground and two scanning disks in the rear.

to the amount of infra-red light incident upon the photo cell. Between the fluorescent lamp and *E*, another 16-in. disk with the same spiral arrangement of holes is spun on the same shaft as the transmitting disk, but 180° out of phase, hence the image is re-inverted. The entire system is enclosed in a light-proof shield provided with a suitable aperture for the projector lens and an observational post pictured in Fig. 2A. Observation of the fluorescent lamp through the spinning disk provides a 32-line replica of the scene and enables *E* to see the discriminanda and animal in sufficient detail to make any desired observation. The arrangement of the spinning disks and lamps can be seen in Fig. 2B. The image definition is suitable to enable *E* to determine which animal is in the situation by looking at the animal's facial characteristics.

A DEVICE FOR RECORDING VARIATIONS IN PRESSURE OF GRIP DURING TRACKING

By WALTER W. SURWILLO, McGill University

The device for recording variations in pressure of grip described here was developed during an investigation dealing with the patterning of muscle-action-potential responses in human tracking. The control unit—the link between *S*'s muscular responses and the tracking system—consists of a shaft which terminates in a concentric wooden knob 3.75 in. in diameter. The knob is so arranged that, with forearm at right angles to the upper arm, a seated *S* can comfortably grasp and rotate it in the manner of a doorknob. Since the rotational axes of the forearm and shaft coincide, appropriate tracking-responses are produced by pronation or supination of the wrist and forearm.

This device, coupled with a resistance-strain gage, furnishes a continuous record of the pressure of *S*'s grip during performance.

The knob (see Fig. 1) was machined from several thicknesses of well-

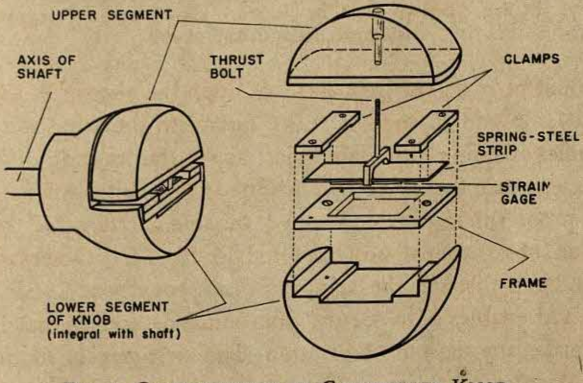


FIG. 1. CONSTRUCTION OF CONTROLLER KNOB

seasoned plywood cemented together; metal fittings were made from ¼-in. aluminum. In assembly, the frame is fastened to the lower portion of the knob which is rigidly secured to the controller-unit shaft. The frame serves as a shelf for the rectangular strip of spring steel as well as furnishing clearance for the strain-gage (Type SR-4) which is cemented to the

* From the Laboratory for Psychological Studies, Allan Memorial Institute of Psychiatry. This work was supported by the Research and Development Division, Office of The Surgeon General, Department of the U.S. Army, under Contract No. DA-49-007-MD-626, and by Grant 405 from the Defence Research Board of Canada.

underside of the strip. Clamps on either side of the strip hold it in place on the frame. One clamp is fastened down firmly over the corresponding end of the steel strip, while the other is so adjusted that the strip can ride smoothly in the channel of the clamp when it is strained. Machining and fitting of the channel are rather critical because with excessive play in it the steel strip wobbles about and causes artifacts in the recording. Too snug a fit, on the other hand, results in hysteresis. The upper segment of the knob is attached to the midpoint of the spring-steel strip by means of the thrust bolt. This is clamped from below to the edges of the strip with a small fitting (not visible in the diagram). The threaded portion of the bolt passes through the vertical counterbored hole in the knob's upper segment where it is made fast with lock-nut and washer.

The knob is so grasped that the thumb makes contact with one segment while the fingers surround the other. An increase in *S*'s grip results in a downward displacement of one segment which is transmitted as a strain to the spring. Variations in strain are recorded as proportional changes in electrical resistance of the strain-gage. Since maximal displacement is usually not more than 1/32 in., *Ss* are rarely aware of the knob's operation.

As shown in Fig. 2, the strain-gage forms one arm of a Wheatstone bridge circuit, the output of which runs to the balanced input of a re-

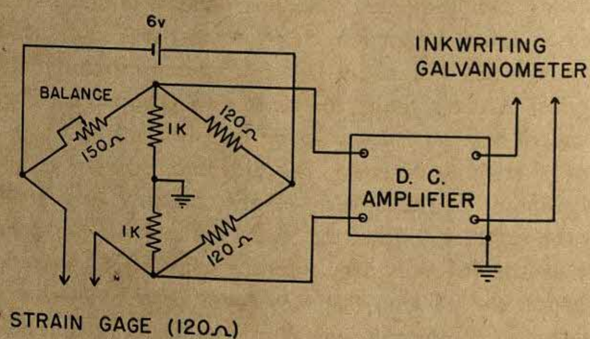


FIG. 2. STRAIN-GAGE BRIDGE AND RECORDING SYSTEM

sistance-coupled amplifier with an amplification of approximately 6,000. Overall sensitivity of the system varies with the dimensions and physical properties of the material strained and the characteristic of the gage. With the present arrangement, a variation of approximately 200 gm. in the pressure of *S*'s grip yields a galvanometer deflection of 1 cm. (1 m.v. input to amplifier). This relation is linear in the range from 50-600 gm.

NOTES AND DISCUSSIONS

ANXIETY AND THE LEARNING OF PAIRED-ASSOCIATES

The relationship between level of drive and performance in a variety of learning situations has been investigated in a series of recent studies, drive being defined in terms of extreme scores on a scale of manifest anxiety.[1] According to the theoretical notions being developed in these investigations, increases in drive should yield superior performance in experimental situations which arouse a single *S-R* tendency, but tend to impair performance in situations which evoke several competing response-tendencies.

Although the major implications of this theory have received experimental support,[2] studies in which the relative superiority of anxious (high-drive) *S*s has been predicted have been confined almost exclusively to nonverbal tasks such as classical defense-conditioning. The evidence here reported concerns the performance of anxious and nonanxious groups in a verbal learning situation in which, according to drive-theory, anxious *S*s should perform at a higher level. Specifically, criterion-groups were tested on two verbal lists, each consisting of eight pairs of nonsense syllables, in which formal intralist-similarity was low. In such low-similarity lists few interfering tendencies were presumed to be present (each stimulus tending to elicit only its own response), and this noncompetitional arrangement was expected to favor the anxious *S*s. A comparison of the anxious and nonanxious groups, each consisting of 17 *S*s scoring extremely high or low on the Manifest Anxiety Scale, confirmed this expectation. The anxious group was superior to the nonanxious on each list both in terms of errors and in terms of trials to criterion. The difference between groups was in each case significant at the 5-% level of confidence or beyond, as determined by the value of *t*. Thus, the results indicate that the superiority of performance of anxious *S*s is not confined exclusively to nonverbal situations and support the position that the amount of response-competition is a relevant factor in determining the interaction between anxiety-level and performance.

Northwestern University — Janet A. Taylor
Neuropsychiatric Institute, Chicago — Jean P. Chapman

[1] J. A. Taylor, A personality scale of manifest anxiety, *J. Abnorm. & Soc. Psychol.*, 48, 1953, 285-290.

[2] *E.g.* I. E. Farber and K. W. Spence, Complex learning and conditioning as a function of anxiety, *J. Exper. Psychol.*, 45, 1953, 120-125; E. K. Montague, The role of anxiety in serial rote learning, *ibid.*, 45, 1953, 91-96; C. K. Ramond, Anxiety and task as determiners of verbal performance, *ibid.*, 46, 1953, 120-124.

DEFOE'S PROJECT FOR LIE-DETECTION

Daniel Defoe (1660-1731) is better known as a novelist and journalist than as a precursor of psychology, but one of his numerous projects for solving the problems of society anticipated by two centuries the scientific approach to lie-detection. In his scheme for preventing street robberies and other disorders of the night, he digressed from his plans for closing the resorts of criminals and discussed the difficulty of dealing with a suspected person who is clever in disguising the usual manifestations of guilt. He wrote as follows:

Guilt carries Fear always about with it; there is a Tremor in the Blood of a Thief, that, if attended to, would effectually discover him; and if charged as a suspicious Fellow, on that Suspicion only I would always feel his Pulse, and I would recommend it to Practice. The innocent Man which knows himself clear and has no Surprise upon him; when they cry, *Stop Thief,* he does not start; or strive to get out of the Way; much less does he tremble and shake, change Countenance or look pale, and less still does he run for it and endeavour to escape.

It is true some are so harden'd in Crime that they will boldly hold their Faces to it, carry it off with an Air of Contempt, and outface even a Pursuer; but take hold of his Wrist and feel his Pulse, there you shall find his Guilt; a fluttering Heart, an unequal Pulse, a sudden Palpitation shall evidently confess he is the Man, in spite of a bold Countenance or a false Tongue: This they cannot conceal; 'tis in vain to counterfeit there; a conscious Heart will discover it self by a faltering Pulse; the greatest Stock of Brass in the Face cannot hide it, or the most firm Resolution of a harden'd Offender conceal and cover it: The Experiment perhaps has not been try'd, and some may think it is not a fair Way, even with a Thief, because 'tis making the Man an Evidence against himself: As for that, I shall not enter into the Enquiry farther than this; if it is agreeable to Justice to apprehend a Man upon Suspicion, if the Particulars are probable and well grounded; it cannot than [sic] be unlawful by any Stratagem that is not injurious in itself, to seek out collateral Grounds of Suspicion, and see how one thing concurs with another.

It may be true, that this Discovery by the Pulsation of the Blood cannot be brought to a Certainty, and therefore it is not to be brought into Evidence; but I insist, if it be duly and skillfully observ'd, it may be brought to be allow'd for a just Addition to other Circumstances, especially if concurring with other just Grounds of Suspicion.

Defoe's proposal differs greatly from the lie-detection procedures of his time, which were based on magical and religious methods and beliefs, and it anticipates the issues, such as its statistical character and its possible invasion of the individual's rights (the Fifth Amendment), which are among those being discussed today.

Indiana University JOHN ROBERT MOORE

[1] Daniel Defoe, *An Effectual Scheme for the Immediate Preventing of Street Robberies and Suppressing All Other Disorders of the Night,* 1730, 34 f. For another case of the anticipation of a modern, psychophysiological discovery, see S. Feldman, Experimental psychology in the Talmud, this JOURNAL, 33, 1922, 304.

SIXTY-THIRD ANNUAL MEETING OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

A few more than 3,500 psychologists and friends registered for the Sixty-Third Annual Meeting of the American Psychological Association in San Francisco, September 1-7, 1955.

A total of 1,011 psychologists read a total of 436 papers and participated in 64 symposia. Twin highlights of the meetings were Lowell Kelly's Presidential Address on "Personality over twenty years," Robert Oppenheimer's address on "Analogies in science." The latter speech, presented on the afternoon of APA Day, was heard by more than 4,000 psychologists and San Franciscans who raised their own hair and probably that of the speaker by joining in a standing ovation for the distinguished physicist. President Kelly's address appears in the November 1955 number of the *American Psychologist.* It is hoped that Dr. Oppenheimer's address can also be made available to psychologists who were not able to attend the meetings.

Many people remarked on the over-all quality of the papers and the symposia. There is clearly an annual increase in the number of psychologists who can and do perform good research and who can and do skillfully present their results in public. Though none can now know for sure, since the progress of science is not highly visible in transit, there surely must have been in San Francisco a significant increment in what Oppenheimer referred to as psychology's "corpus of certitude."

Theodore M. Newcomb was inducted as President of the Association and Lee J. Cronbach as President-elect. Neal E. Miller and Dorothy C. Adkins were elected to three-year terms and Donald W. MacKinnon to a one-year term on the Board of Directors. Launor F. Carter will serve as the Association's Recording Secretary. Arthur W. Melton and Harry F. Harlow were elected to second terms as editors of the *Journal of Experimental Psychology* and the *Journal of Comparative and Physiological Psychology* respectively. Actions of the Council of Representatives included revisions, required by postal laws and financial necessity, in subscription prices to APA journals, the establishment of three $1000 APA awards for outstanding scientific contributions. The election of 1,269 new Associates and 124 Fellows brings to 14,600 the total membership of the Association and to 1,447 the total number of Fellows.

The 1956 meetings of the Association are to be held in Chicago, Illinois.

Washington, D.C. FILLMORE H. SANFORD

TWELFTH CONGRESS OF THE ASSOCIATION INTERNATIONALE DE PSYCHOTECHNIQUE

The XIIth Congress of the Association Internationale de Psychotechnique was held in London during the week of July 18-23, 1955. More than 600 members and others, representing 35 countries were in attendance. Representatives from the United States totaled approximately 50 in number.

A significant feature of the meeting was action in changing the name of the organization from *Association Internationale de Psychotechnique* to *Association Internationale de Psychologie Appliquée.* This move reflects a growing recognition of the research foundations of applied psychology, as contrasted with the earlier tendency—particularly on the part of Europeans—to view applied psychology largely as a system of techniques and procedures. Involved also is the desire to relegate industrial applications, which gave rise to the term "psychotechniques," to an appropriate position as only one of the many areas of applied work with which the psychologist is concerned.

This change in orientation was also apparent in the make-up of the program which included symposia on the "Social psychology of the classroom," "Problems of ageing," "The psychologist and society," "Communications problems," "Attitudes," "Opinions and behavior," along with a number dealing with specifically industrial applications of psychology. A small number of individual papers were presented, but the program was devoted predominantly to symposia and lectures arranged by invitation. Among the lectures was a discussion of "Timing—a fundamental character in human skill," by Sir Frederick Bartlett (England); "Influence of foresight and attitude on mental output," by Professor E. Mira y Lopez (Brazil); and a closing address entitled "The new Utopia," by Professor Morris S. Viteles (U.S.A.).

An interesting feature of the meeting was the use of simultaneous multilingual translation made possible by a grant from the European Productivity Agency. A series of receptions and a banquet addressed by the Earl of Halsbury, Chairman of the National Institute of Industrial Psychology, contributed to the success of an outstanding series of meetings.

University of Pennsylvania MORRIS S. VITELES

Stephen Polyak: 1889-1955

Dr. Stephen Polyak was found dead on the morning of March 9, 1955. Fate decreed that he, the outstanding authority on the anatomy and pathology of the visual system, should be the victim of disease processes that in the course of years gradually and almost entirely destroyed his extramacular vision. It is remarkable with what energy he resisted the encroachments of this disease: even the day before his death he devoted to work on his last book. The writing of this book, a comprehensive volume on the vertebrate visual system, was, indeed, a heroic accomplishment achieved in a period of ill health during the last ten years. It is true that it could not have been accomplished without the never-failing support of his wife and, it should be added, not without the help of the University of Chicago which supported his work and made it possible for him to continue his investigations.

The story of his scientific life can be succinctly told by giving the complete titles of the three books which he wrote in the course of his career and which embody the results of his scientific endeavors: (1) *The Main Afferent Fiber Systems of the Cerebral Cortex in Primates: an Investigation of the Central Portions of the Somato-sensory, Auditory, and Visual Paths of the Cerebral Cortex, with Consideration of Their Normal and Pathological Function, Based on Experiments with Monkeys* (University of California Press, 1932). (2) *The Retina: the Anatomy and the Histology of the Retina in Man, Ape, and Monkey, including the Consideration of Visual Functions, the History of Physiological Optics, and the Histological Laboratory Technique* (University of Chicago Press, 1941). (3) *The Vertebrate Visual System: its Origin, Structure, and Function and its Manifestations in Disease with an Analysis of its Rôle in the Life of Animals and in the Origin of Man.* Preceded by a historical review of investigations of the eye, and of the visual pathways and centers of the brain. (In press, University of Chicago Press, a royal octavo volume of about 1600 pages, including 551 illustrations, 27 in color, and a bibliography of about 10,000 references). As he has pointed out in the preface of the last book, this work represents the outcome of "the almost continuous labor of nearly three decades." He himself read the galley proofs of the first fourteen chapters and it is hoped that this monumental volume will be published in the not-too-distant future.

Born on December 13, 1889, in northern Croatia (Yugoslavia), Stephen Polyak was educated at the Classical Gymnasium in Zagreb and at the University of Graz, Austria. His medical training that he began in 1909 at the Medical School in Graz was interrupted in 1914 by the events of World War I. While attached to the Medical Corps of the Austro-Hungarian army, he participated in the battles at the Russian front during the first three months of the war, but then was captured by the Russians and interned at Kiev. Subsequently he was employed as a physician in a Russian military hospital and in the Medical Corps of the Serb Volunteer Divisions. Although the General Staff of the Russian army granted his request to return to Serbia and issued a special passport to him, he was unsuccessful in his attempt to reach Serbia. He was, however, permitted to register at the Medical School of the New-Russia University in Odessa where he received, after studying medicine for two semesters, his M.D. degree in the fall of 1916. It is worth recording that he saw, during a review of the Serb troops in Odessa, the Czar and his family at close range, little realizing, as he wrote later, that they would be murdered in about a year and that he was in fact witnessing "the funeral cortège of the last Romanovs . . . and the last pale glory of Russia's Empire on that warm summer day in the year 1916."

The 1917 revolution found him on the medical staff of the Second Serb Volunteer Division, in the divisional field hospital located in a small town of southern Russia not far from the Dnieper river. With other members of the Serb divisions he was permitted to leave Russia by way of Archangel on the White Sea. The Russian ship that took him around Norway to Scotland anchored at the entrance of the Cromarty Firth at the beginning of October 1917. His subsequent journey took him through England and France (where he almost died as the result of an inoculation by a French sanitary soldier) to Salonika in Greece, where he joined the Serbian army and participated in its battles in Macedonia and Albania until the end of the war.

It was only in 1920, after about six years in uniform, that he became a civilian again. In the same year, the University of Zagreb granted him an M.D. degree. While an assistant in the Department of Neurology and Psychiatry in this university (1920-28), he became actively engaged in research in Zagreb and also in other centers. A letter of his, written in French and addressed to the Instituto Cajal in Madrid, brought him to his surprise a complete set of the "Trabajos" of the University of Madrid which he diligently studied during the following years.

His researches took him in 1921 to the Neurological Institute at the University of Vienna (Obersteiner's Institute), in 1924-25 to the University of London (for work with G. Elliot Smith), in 1925 to Madrid (to visit Cajal), and in 1926-27 to Chicago where he was attracted by the work of C. Judson Herrick and K. S. Lashley. On returning from Chicago to Europe, he found conditions at the University of Zagreb and in Yugoslavia such that he decided in 1928 to leave his native country and accept an offer from the University of California. There he served as Assistant Professor of Neuroanatomy until 1930 when he returned to Chicago to spend the remaining twenty-five years of his life. From 1930 to 1937 he was Assistant Professor and then Associate Professor of Neurology in the University of Chicago Clinics. From 1937 on he was a member of the Department of Anatomy, in which he first held an appointment as Associate Professor until 1942, and then as Professor of Anatomy from 1942 until his death.

Pavlov's advice as quoted by William Bragg was superfluous for Stephen Polyak: "Remember that science demands from a man all his life. If you had two lives that would be not enough for you. Be passionate in your work and your searchings." Stephen Polyak certainly belonged to those scientists who in the Pavlovian sense are 'passionate' in their work and their searchings. Ever since the guns turned silent on the European battlefields at the end of the First World War, he worked almost uninterruptedly and, it may be said, with a burning intensity. His interests extended over a wide domain, from biology and natural history to anthropology and politics, from the history of science to the history of nations and of Western civilization itself.

It is surprising that he found the energy and the time even during the last months of his life to put in order and work on a manuscript of about 2000 typewritten pages, a fourth book, a diary dealing with the story of his life and our generation. This book, entitled *Glory to Them All,* is another testimony to the universality of his interests; and it is characteristic that it has the subtitle *Recollections of a Nobody.*

Stephen Polyak's life, unlike that of many other scientists, was not marred by the seeking for fame, power, advancement, and 'the glories of this world.' No wonder, therefore, that the world did not seek him out to heap special honors upon him as the years sped on. He was so intensely absorbed in his work that he did not trouble to join many societies or attend many scientific meetings, although he consented to serve on the editorial board of the *Journal of Comparative Neurology* from 1937-48.

The tributes paid to him by leading neuroanatomists and other scientists, however, leave little doubt that his work will stand the test of time longer than will generally be the case with the contributions furnished by modern scientists. His fundamental contributions, it must also be remembered, were made almost single-handedly. Those who were able to witness his methods of work through the years and his attitude towards scientific pursuits could not help but feel that Balfour's characterization of Darwin was also applicable to Stephen Polyak: "He was not a partizan—he really wanted to find out the truth—an attitude of mind seldom found among men of science." His high standards of scientific integrity, the honesty and meticulous care with which he worked or drew conclusions from his work, his humility in the face of intricate and unsolved problems—these virtues as found in Stephen Polyak do not, it is unfortunately true, readily grow or thrive in the microclimate of present-day scientific institutes or universities. Being devoid of personal ambition, he lived a life in which ultimately nothing mattered except his scientific work, and the problems, demands, and exigencies arising in connection with this work. Those words of Francis Bacon that have subsequently been used to introduce one of the most famous books ever produced in the history of Western thought appear to be the very words written across all of Stephen Polyak's scientific endeavors: *De nobis ipsis silemus. . . .*

University of Chicago HEINRICH KLÜVER

Walter Fenno Dearborn: 1878-1955

Walter Fenno Dearborn died in St. Petersburg, Florida, the morning of June 21, 1955, after a crippling illness which followed a severe cerebral hemorrhage on November 14, 1953. He was born in Marblehead, Massachusetts, July 19, 1878.

Dearborn attended the Boston Public Schools and graduated from Phillips Exeter Academy in 1896. He received an A.B. degree from Wesleyan in 1900 and an A.M. in 1903. To finance his schooling he occupied the positions of vice-principal and teacher of Latin and Greek at the Middletown, Connecticut, High School for three years. In 1903 he went for further graduate work to Columbia, where he received a Ph.D. in educational psychology in 1905. He studied at Göttingen in 1904, at Heidelberg in 1911, and received an M.D. degree from Munich in 1913. Harvard gave him an honorary A.M. in 1941.

Immediately after graduating from Columbia he went to Wisconsin as an instructor. There he was promoted to an assistant professorship in 1907. In 1909 he left to become associate professor of education at Chicago, being presently enticed away from Chicago by Harvard in 1912 with the offer of an assistant professorship and the promise of a professorship in the near future. In 1917 the promise was made good; he was promoted to a professorship and made director of the Psychological Clinic of the Harvard Graduate School of Education. He was retired as emeritus professor in 1947 and became professor of psychology and education and director of the Psycho-Educational Clinic (a laboratory for reading problems) at Lesley College, Cambridge, Massachusetts, a position he occupied until he became incapacitated.

Dearborn was a trustee of the Massachusetts State Infirmary in 1913–14 and 1915–33. He was also a trustee of the Walter E. Fernald School from 1942. He was a member of the Council of the American Psychological Association from 1918 to 1920, a Fellow of the American Academy of Arts and Sciences, Associate Fellow of the Massachusetts Medical Society, and an honorary member of the Hungarian Society of Education.

An educational psychologist whose time was primarily devoted to research, Dearborn centered his interest on three subjects—reading problems, the relation of physical growth to intelligence, and intelligence tests—and in each area he plowed new ground. He and his many devoted students published numerous papers on these subjects, many of which appeared in

the *Harvard Monographs in Education* under the general title of "Studies in educational psychology and eductaional measurement." This series was edited by Dearborn from 1922 to 1926.

It is significant that Dearborn's doctoral thesis was "The psychology of reading: an experimental study of the reading pauses and movements of the eye,"[1] for this first interest remained his most absorbing one and it was in this field that he made his greatest contribution to psychology. In a monograph entitled "Special disabilities in learning to read and write,"[2] in which he coöperated with E. E. Lord and Leonard Carmichael, he contributed a section on "The etiology of congenital word blindness." He called special attention to the fact that lack of motivation seemed a very strong factor in disabilities in reading and writing, a fact that had not up to that time been sufficiently emphasized. In 1939 he found that there were part-word, whole-word, and mixed-type readers.[3] Two years later, together with I. H. Anderson, he showed that anesekonia is one of the many factors contributing to disability in reading.[4] He and Anderson also found that there is a positive relation between reading ability and college achievement even when only those differences are considered which are independent of intelligence.[5] In 1947 he and Carmichael published under the title of "Reading and visual fatigue" the results of their long and careful series of experiments in which students were required to read continuously for 6-hr. periods.[6] They found that this extreme effort resulted in no cost to the organism. In 1952 appeared *The Psychology of Teaching Reading*,[7] based in great part upon the previous work of the Dearborn group. The point was stressed that in teaching reading motivational, emotional and cognitive variables must be considered.

The problem of the relation of anatomical age to intelligence occupied Dearborn and his students for many years. He directed an investigation of individual development under a grant from the Commonwealth Fund in which radiograms of the stages of ossification, particularly of the carpal bones of children, were examined. More exact objective measurements

---

[1] *Op. cit., Columbia Contr. Philos., Psychol. & Scient. Methods*, 4, 1906, 1-134.

[2] E. E. Lord, Leonard Carmichael, and W. F. Dearborn, *op. cit., Harvard Monog. Educ.*, 2, 1925, 1-76.

[3] Dearborn, The use of the tachistoscope in diagnostic and remedial reading, *Psychol. Monog.*, 47, 1936 (No. 212), 1-19.

[4] I. H. Anderson and W. F. Dearborn, Anesekonia as related to disability in reading, *J. Exper. Psychol.*, 23, 1928, 559-577.

[5] Anderson and Dearborn, Reading ability as related to college achievement, *J. Psychol.*, 11, 1941, 387-396.

[6] Dearborn and Carmichael, *op. cit.*, 1947, 1-483.

[7] Anderson and Dearborn, *op. cit.*, 1952, 1-382.

were made in this than in previous studies where the estimates were mostly subjective. In his "The mental and physical development of public school children" Dearborn reviewed the findings of a very extensive study of the mental and physical growth of 3000 children over a period of 12 yr.[8] This work was the forerunner of Harvard's Laboratory of Human Development. In 1941 Dearborn and J. W. M. Rothney published their book *Predicting the Child's Development,*[9] in which they analyzed the data from the Harvard Growth Study, begun in 1922 by the longitudinal method. Dearborn came to the final important conclusion that "physical and mental growth are essentially individual affairs" and "the relation between the physical measurements and the mental measurements is so low that the knowledge of one does not enable us to predict the other."[10]

Dearborn's greatest contribution in the field of tests is probably *The Dearborn Group Tests of Intelligence,* which appeared in 1920.[11] He also published in 1923, together with E. A. Shaw and E. A. Lincoln, "A series of form board and performance tests of intelligence."[12] He could not decide whether performance tests are a measure of general intelligence, but he found them very useful in clinical examinations. Dearborn gave a series of eight lectures at the Lowell Institute and he published a revised and amplified account of these lectures in 1928 under the title "Intelligence tests: their significance for school and society."[13] He included a section on poor readers and emphashized the influence of environment as against heredity in regard to delinquency.

Walter Dearborn was a quiet, sensitive, taciturn man with a strong drive and great perseverance. His mildness was combined with strong convictions and a stubborn support of the best values of his profession. He was careful and intelligently critical in his research. His contributions were sound and have been of great benefit to the development of the special fields in which he was interested.

Princeton University HERBERT S. LANGFELD

[8] *Op. cit., Sch. & Soc.,* 41, 1953, 585-593.
[9] *Op. cit.,* 1941, 1-360.
[10] *Idem,* 341.
[11] *Op. cit.,* 1920, 1-28.
[12] E. A. Shaw, E. A. Lincoln, and W. F. Dearborn, *op. cit., Harvard Monog. Educ.,* 1, 1923, 1-64.
[13] *Op. cit.,* 1928, 1-360.

Walter Dill Scott: 1869-1955

Walter Dill Scott, born on May 1, 1869, and President of the American Psychological Association in 1919, was professor of psychology at Northwestern University from 1901 to 1920 and president of the University from then on until his retirement in 1939 at the age of seventy years. He received his A.B. from Northwestern in 1895 and his Ph.D. from Leipsig in 1900. He died on September 24, 1955, at the age of eighty-six years.

Scott was an applied psychologist, interested in solving practical problems rather than in developing psychological theory. He may properly be called the father of applied psychology for no one else applied psychology to such a variety of business problems as he did and at so early a date.

His earliest applications of psychology pertained to advertising. His two books, published in 1903 and 1908, preceded all others except that of Harlow Gale, whose work was privately published in 1900. He was also interested in the application of psychology to selling, publishing *Psychology of Public Speaking* in 1907 and *Influencing Men in Business: The Psychology of Argument and Suggestion* in 1911.

On leave of absence from Northwestern University, Scott spent the year 1916–1917 at Carnegie Institute of Technology as director of the Bureau of Salesmanship Research. Five aids to the selection of salesmen were formulated then; one of them was a rating scale. After war was declared, he adapted this scale for use in selecting officers. With the coöperation of Edward L. Thorndike and Frederick P. Keppel, then assistant to the Secretary of War, he persuaded commandants in several officer training camps to use his rating scale, and immediately afterward he sold the idea to Army officials in Washington. His efforts led to the first widespread use of rating scales.

One special project assigned to Scott was the rating of all officers of the rank of colonel and higher with the exception of Pershing. Each was rated by ten outstanding generals and the average given to the Secretary of War. After the war it was found that nine of the ten with highest average had received special honors and promotions and all ten of lowest average had been sent home.

The Secretary of War, Newton D. Baker, was very desirous that enlisted men should be assigned as far as possible to work for which they were best fitted. At that time there was no mention of classification and assignment of personnel in the revised Army staff annual. After several conferences with Scott, General H. P. McCain (the Adjutant General), and several others, Baker appointed the Committee on Classification of Personnel in the Army

to accomplish his purpose. Scott served as director, Thorndike as chairman and Walter V. Bingham as secretary, with seven other psychologists and two industrial personnel men. This was a unique organization—a committee of civilians within the framework of the Army, which determined policies, introduced necessary procedures and supervised their execution. The committee not only elaborated procedures already recognized in business but introduced new ones, as for example, trade tests, and the ascertainment of the supply of skilled workmen within the country and demand for them within the armed forces.

In recognition of his service to the country, Scott was awarded the Distinguished Service Medal for "devising, installing and supervising the personnel system in the U. S. Army." He continued his interest in this field as joint author of *Personnel Management* (1923), the fourth revision of which appeared in 1949.

Scott was a leader, not a domineering man who gave orders but one who made it easy and logical for others to follow him. He was a great salesman, not of commodities but of ideas. He not only wrote a book on suggestion in selling, but used his theories—more than most text-book writers do. He was master of *reductio ad absurdum.* He would open an interview with an officer by saying, "Let's see if I have your idea clearly in mind." Step by step he would lead to a conclusion to which the officer would eventually exclaim, "But that's all wrong." Then they would start all over with Scott's solution.

Scott believed that everything possible should be done by him to make it easy for another to say *yes.* He spent nearly a week preparing a letter that looked exactly as though the Secretary of War had dictated it asking the President for half a million dollars for the personnel program. The time, from the moment Scott had said "Good Morning" and handed the letter to Baker until the latter had signed it, was 55 seconds according to his stopwatch. In his later days he had ready admittance to busy officials since they knew he would spend only a minute or two of their time.

Scott's record of raising $70,000,000 for Northwestern University during his nineteen years as president is another example of his salesmanship.

Scott was a lovable character, perceptive and kind. He was always mindful of the feelings of others, freely expressing approval of their behavior. His criticisms were privately expressed to associates when seeking advice as to how the situation could be improved. He engendered morale, keeping always before his associates the bigness, complexity and worth-whileness of the task and arousing their enthusiasm for its accomplishment.

Those who knew him will not forget him.

Stanford University EDWARD K. STRONG, JR.

BOOK REVIEWS

Edited by M. E. BITTERMAN, The Institute for Advanced Study

Le Méchanisme de la Vision des Couleurs: Physiologie—Pathologie. By J. SÉGAL. Paris, G. Doin, 1953. Pp. 351.

This well-organized and provocative monograph, with its revolutionary theoretical proposals, is based on a number of experimental studies by the author, and backed by citations, curves, and tables from well known workers in the fields of retinal histology, photochemistry, pathology, and by psychophysiological observation. Among the 125 research workers cited, covering the last decade in Europe especially, are Polyak, Le Gros Clarke, Lythgoe, Wald, Granit, Willmer, Hartridge, Hartline, Pitt, Wright, Purdy, Hecht, and others. Ségal's own research was carried on in the Laboratory of Sensory Physiology of the Collège de France, founded by Henri Piéron, whose preface to the volume commends the originality and industry of his former student and certain of the positions taken in the text (*e.g.* as to the ambiguity of the Rayleigh equation) though reserving judgment on certain others.

The first six chapters outline Ségal's thesis. Recognizing that trichromatic theory —resting on three hypothetical photochemicals, selectively responsive to red, green and blue radiations, and located in a mosaic of cone receptors—has reached a stalemate, he proposes to reject all suppositions, for which no proof has been forthcoming in a century and a half, and start afresh with the facts in hand. The one known retinal photosensitive substance is visual purple or rhodopsin. (Wald's iodopsin of 1938 is rejected as an artifact, while this work antedates Wald's later findings.) This, with its derivatives and antecedents, Ségal proposes to make do double duty, in dark and daylight.

Working within the framework of trichromatic theory, that is, accepting the fact that for the normal-visioned any hue in the spectrum may be reproduced by a mixture of three wavelengths—provided the three occasionally are shifted—Ségal proposes a new working hypothesis as to receptor mechanisms. Starting from recent chemical and histological studies and analyses, he proposes rhodopsin in three different phases, and three different locations, as the possible triple substrate. Macular pigment (identified with Lythgoe's 'transitory orange,' a phase in rhodopsin evolution or regeneration, and located in the fibres of Henle near the inner surface of the retina that adjoins the vitreous humor) is set up as the substrate of blue or violet. Rhodopsin in liquid form, in the tips of both rods and cones, must serve for green; while rhodopsin in solid or crystalline form, in the pigment epithelium, between the tips of rods and cones and the choroid, will pick up the light waves not already filtered out by the violet and green receptors and give rise to the quality of red.

These postulates are backed in the text by numerous absorption curves and cuts of retinal cross-sections. The hypothesis is christened that of "les trois couches," or *the three layers,* and is frankly launched not as demonstrated fact, but as a stimulus to fresh experimentation. The reader, at this point assailed by doubts as to

the problem of visual acuity or configurational definition, with red and green receptors crowded off the rod-and-cone proscenium, is referred to the six succeeding chapters for enlightenment and evidence for or against the thesis. To anticipate, the acuity problem is side-stepped by assigning the green receptors (cones) the entire responsibility for acuity, though Piéron's argument for a separate "white cone," and Granit's achromatic "dominators," are respectfully noted.

Verification of the three-layer postulates is sought first in experiments in chromatic adaptation with stimuli of extraordinarily high intensity—12,000 to 7,000,000 trolands—and secondly, in case histories of dichromatopsy, congenital and acquired, here designated as 'chromatic anomalies.' The fact that after intensive stimulation the red quality is the first to reappear in the spectrum is accepted as proof that the red receptor is located in the chromatic epithelium, which lies next to the choroid, the layer responsible for regeneration of rhodopsin. A colored plate (of doubtful accuracy in the opinion of the reviewer) depicts the recovery of hues throughout the spectrum following different intensities of stimulation and also after exposure to red-free, green-free, or blue-free lights (p. 200, six different modified spectra and one normal shown, Ilford and Wratten filters used in three cases).

In Chapter XI, the problem of daltonism or protanopia—a stumbling-block to every theorist—is solved with startling boldness. The dyschromatopsies are all explicable, not by absence of a given type of receptor or photochemical substance, but by diminution or augmentation of the amounts of the three different phases in their three respective layers. Deuteranopia (once known as green-blindness) is caused by an excess of rhodopsin in the rods and cones, and hence excess of the green sensation in the image; protanopia, by excess of the crystalline form in the epithelium—leaving the red sensation intact and enhanced, though the long-wave end of the spectrum is shortened (owing presumably to over-absorption in the cones). Counterbalancing, the scepticism aroused by these paradoxial conclusions is a valid criticism echoed today in many quarters, directed against the customary interpretation of the Rayleigh equation as obtained with the Nagel anomaloscope setting. The preponderance of red in the *R/G* mixture needed to match a certain yellow indicates, according to Ségal, deficiency not in the *red* sensory substrate but in the *green*—owing perhaps to a shift of the yellow tonality toward red.

The 'tritanopia' of the central fovea, first reported by König and rediscovered by Willmer in 1946, is cited as proof of the absence of blue receptors, replaced by green. The fact is overlooked, here as elsewhere, that blue and yellow are both high-threshold colors, and that in the congested foveola there is no chance for the summation of stimuli through the aid of mop or brush bipolars as in eccentric areas. In classifying *S*s as protans and deutans, by the way, Ishihara plates are supplemented by Polack's sorting test (1939), with its many variants of hue, chroma, and luminance.

While the arguments advanced for *les trois couches* may leave the reader cold, the scientific service rendered by Ségal in emphasizing the importance of continued investigation of the chemical properties of rhodopsin and related substances, along with his emphasis on recent findings in European laboratories, must be acknowledged. His recognition also of the complexity of deviations from normal vision and need of further study, his boldness in throwing overboard assumptions that have proved useless, his ingenuity in devising new experimental approaches and more fertile hypotheses, combine to render this a unique contribution to color science.

The volume deserves careful study by specialists in many branches involving color problems, if only to discover the flaws in the reasoning. This holds for ophthalmologist, physiologist, and illuminating engineer, as well as psychologist concerned with the actual problems of perception. The careful perusal in detail of Ségal's arguments will serve as a useful shakeup—an inventory and reassessment—challenging the reader's reasoned knowledge, supported by experimental proof, on many points of theoretical and practical importance in the field of color perception.

Cornell University ELSIE MURRAY

The Achievement Motive. By DAVID C. MCCLELLAND, JOHN W. ATKINSON, RUSSELL A. CLARK, and EDGAR L. LOWELL. New York, Appleton-Century-Crofts, 1953. Pp. xxii, 384.

This book makes three more-or-less discrete contributions to the study of motivation. The first contribution is a theory of motivation; the second large section carefully describes the measurement of the achievement motive through content-analysis of imaginative stories; the rest of the book summarizes a number of experiments with the achievement motive as the dependent variable.

In an essay labeled "Toward a Theory of Motivation," the essential features of contemporary thinking about motivation are reviewed. The authors summarize three competing motivational models and find them wanting. Less acceptable than the preferred *affective-arousal model* are the survival-model (Hull, Freud), the stimulus-intensity model (Dollard and Miller), and the stimulus-pattern model (Hebb). There is little difficulty in rejecting the survival model with its central theme of tissue-needs tripping off need-reducing behavior. A large number of tissue needs may be created, as in avitaminoses, without the organism behaving in ways which are normally regarded as driven or motivated, and there are experiments which suggest that instrumental learning is produced directly by afferent stimulation rather than by slower-developing tissue changes. The authors also make short shrift of the stimulus-intensity model, for which the best that can be said is that strong stimuli *may* be one of many motivating conditions. The stimulus-pattern model is rejected primarily because it is too general. The authors argue that Hebb's phase sequences and the like are of little value to the experimentalist because they cannot be specified beforehand. How, for example, can the discrepancy between expectancy and perception in human conduct be assessed when the expectancy cannot be specified in univocal terms?

The affective-arousal model espoused by the authors is really an extension of the stimulus-pattern model. A motive is defined as "the redintegration by a cue of change in an affective situation" (p. 28). The core of the theory is expressed as follows: "Certain stimuli or situations involving discrepancies between expectation (adaptation-level) and perception are sources of primary, unlearned affect, either positive or negative in nature. Cues which are paired with these affective states, changes in these affective states, and the conditions producing them become capable of redintegrating a state (A′) derived from the original affective situation (A), but not identical with it. To give a simple example, this means that if a buzzer is associated with eating saccharine the buzzer will in time attain the power to evoke a motive or redintegrate a state involving positive affective change. Likewise, the buzzer if associated with shock will achieve the power to redintegrate a negative affective state. These redintegrated states, which might be called respectively

appetite and *anxiety,* are based on the primary affective situation but are not identical with it" (p. 28).

How do we know when affective arousal occurs? The authors answer that physiological measures of autonomic changes are satisfactory but by no means ideal indicators. Autonomic measures cannot, however, distinguish between positive and negative affect, between appetite and anxiety. With humans, the nature of an affective state can be assessed by such indicators as expressive movements, impromptu vocalizations, and, of course, fallible verbal statements. With non-verbal *Ss*, the kind of affect may be inferred from innate response-patterns and from approach and avoidance behavior.

A section on "antecedent conditions for affective arousal" is a scholarly and challenging bit of scientific writing. Data and argument are adroitly introduced to support such propositions as the following: that the origin of affective arousal is innate; that positive affect results from small discrepancies between perception and adaptation level, while negative affect results from larger discrepancies; that departures from expectancies result in positive-negative affect in *either* direction; that events can differ from expectancies on a number of dimensions; and that multiple and complex human adaptation levels may endow a single event with a multiplicity of affective properties.

In this essay, the authors restructure the persistent problem of primary and secondary drives. In their view, all motives are acquired, but the affective arousal upon which motives are built is innate. The traditional distinction between biological (primary) and social (secondary) drives is abandoned. The arguments are convincing and open the way to a fresh attack on motivational problems.

Because much of the theoretical portion of the book was written after the experimental work was completed, the connections between theory and the studies of "thought samples" from imaginative productions are tenuous. The reviewer is not convinced that the achievement motive as defined necessarily follows from the theory. Expectations, which precede the arousal of the achievement motive, are described as "standards of excellence." The motive is aroused when discrepancies from such expectations produce positive or negative affect. In the *TAT* stories, the expectations are inferred from statements involving the *evaluation* of performance; the affect is inferred from somewhat arbitrarily selected contents of the stories, such as "he is *unhappy*" or "he *wants* to be successful." The theoretical arguments led the reviewer to expect that adaptation level would be assessed in a fairly rigorous way and that affective arousal would be determined by objective, autonomic measures. These expectations were not confirmed and the resulting affect was negative—a phenomenological report which provides support for at least part of the theory. Inferences from imaginative productions may or may not represent levels of adaptation, and third-person stories may or may not reflect affective arousal.

The second contribution of this volume is devoted to a careful description of the method of obtaining "thought samples," a method which the authors recognize to be quite different from the clinical *TAT*. The *Ss* are tested in groups, under strict time-limits, and are asked to write imaginative stories to cover four specific questions. The procedures for analysis of content are clearly drawn. The several indices are summed to give a single measure of the achievement motive (as defined).

The third part of the volume analyzes a score of empirical studies by McClelland

and his collaborators, some of which are reported elsewhere. Among the studies are several that relate achievement-imagery to experimental sets established just prior to the "thought sampling." As expected, the prior activity influences the story-writing. In an attempt at functional validation, the authors show that the achievement-imagery is related to a number of independent (and sometimes miniscule) variables, among them performance in simple arithmetic under speed-pressures, improvement in the ability to solve anagrams, recall of incompleted tasks, college grades, and the like. The data which are used to show that students who score high on achievement-imagery earn higher college grades are somewhat ambiguous. Whether or not the obtained correlations may be considered as support for the inference that scores in achievement-imagery reflect the achievement motive is not convincingly established. Judgments by clinicians of the presence of the achievement motive have little relationship with the achievement-imagery scores. The reviewer, like the authors, would be happy to see the imagery index validated against other effects of the achievement motive drawn from life-situations. If such further validation holds up, then McClelland and his co-workers are to be congratulated for having isolated a measurable social motive without recourse to real or imagined physiological antecedents.

It is difficult to evaluate the book as a whole. The essay on motivational theory stands out as a fine piece of critical writing, but the fruitfulness of the effective-arousal model remains to be established. The sections of the book that deal with the development of the procedures for content-analysis are more in the nature of a test-manual and must be evaluated as a good job. Insofar as the experimental studies are concerned, it is the reviewer's judgment that they are better suited to journal publication, although a book, to be sure, provides more freedom for discussion and speculation. As the authors modestly say, they have given us a report of research in progress.

University of California THEODORE R. SARBIN

Theory and Problems of Adolescent Development. By DAVID P. AUSUBEL. New York, Grune & Stratton, 1954. Pp. xviii, 580.

Psychology of Adolescence. By LUELLA COLE. Fourth Edition. New York, Rinehart and Company, 1954. Pp. xvi, 712.

In many ways these two books are a study in contrast. Cole's *Psychology of Adolescence* is the fourth edition of a standard undergraduate text. It lists 151 illustrations and 59 tables, and is enlivened by numerous case histories and personal anecdotes. Ausubel's *Theory and Problems of Adolescent Development* (he acknowledges his debt to Krech and Crutchfield for the title, and to Sherif and Cantril for the overall theoretical orientation) is described in the preface as "primarily intended as an advanced textbook in adolescent psychology for graduate students in psychology and education" (p. xvi). It contains only three illustrations and three tables. "The criterion followed in this matter was that graphic and tabular material would be included only if the concepts to be communicated could *not* be more succinctly and precisely conveyed by words. . . . The resulting space that was saved was used for integrative and interpretative purposes" (p. xv). Case histories were excluded on the ground, probably correct, that they are too often accepted as evidence rather than illustration. Cole continues the (probably sound, but sometimes distracting) custom of giving details of the samples and results of the various investigations cited. Ausubel eliminates them entirely, covering himself with a prefatory caution

and an admonition to go to the sources. Cole devotes almost 200 pages to mental development and to intellectual and vocational interests. Ausubel covers the same area in about a hundred pages.

Cole's text is characteristic of many which deal with developmental psychology. The plethora of empirical material sometimes overwhelms the reader without providing much in the way of a conceptual mooring. Ausubel asserts that "logically prior to obtaining definite evidence in the field of adolescent psychology is the necessity for relating research efforts to integrative hypotheses derived from a comprehensive theoretical orientation to the nature of adolescent development" (p. 136).

Ausubel's view is that "adolescence is a distinctive period of personality development that includes certain unique and universal maturational changes requiring extensive reorganization of personality structure" (p. 148), and he makes a determined effort to show how the well-known characteristics of the adolescent period can be better understood from this point of view. Such factors as pubescence, social expectations of independence in some areas, but parental demands for submission in others, "transitional anxiety," somatic variations, and the decreasing importance of "derived" status are all tied to the central problem of personality change. Thus, in a discussion of the adolescent peer culture Ausubel observes that it is facilitated by the adolescent's greater mobility and that it provides for the gratification of newly developing heterosexual needs. "Much more important, however, is the adolescent's increasing concern with acquiring primary status as an independent entity. And since there can be no status apart from a system of relationships to a constituted social unit, adolescent peer groups are 'formed spontaneously to serve the function of a social institution, to secure a status and a social identity for youngsters not genuinely provided with such an identity by society at large' " (pp. 342-343).

In short, the difference between these books is the difference between an ambitious attempt at a comprehensive theoretical organization, and a modest up-to-date revision of a traditional and fact-filled text. In the reviewer's opinion Ausubel does not quite succeed. His theory is less a theory than a point of view, and in his effort to make it a theory he has substituted assertion and reiteration for close reasoning. The incessant quotation of his own previously published opinions gives the impression that he has tried and failed to find other support. It also leads an unwary reader to assume that the quotations are from others who have thought about the problem, and to draw the conclusion that Ausubel *has* that support. Ultimately the reader comes to resent the quotations, to look for the airy assertion, and to wonder whether the somewhat prolix style might not have profited from a judiciously administered blue pencil. The space saved could then have been devoted to concrete detail, graphical and tabular material, and readers could try to draw their own conclusions. Cole, having attempted a less grandiose project, comes much closer to her mark. Except for a treatment of sexual behavior and motivation that is much too brief, hers is a thorough textbook.

There remains one additional comment about these two books that the reviewer cannot forbear, even though it pertains to the marts of trade rather than the realm of academe. Cole's 700 pages are profusely illustrated, include many graphs and tables, are well bound, and cost six dollars; Ausubel's 580 pages have almost no illustrations, graphs, or tables, are poorly bound, and cost ten dollars. This bit of publishing banditry should not go uncensured.

University of California JOHN P. McKEE

Mind and Performance: A Comparative Study of Learning in Mammals, Birds, and Reptiles. By HAROLD K. FINK. New York, Vantage, 1954. Pp. xi, 113.

Here is a book which attempts to compare the learning ability of turtle, tortoise, chicken, rat, cat, dog, pig, and man, but the methodological, statistical, and logical inadequacies of the work make it valuable only as a 'what not to do' example for introductory courses. A maze was used, which consisted of four alleys emanating from a central choice-point. After one alley was learned to criterion, another was designated as correct, and the animal was required to extinguish the formerly successful response, acquire a new alley-response, and so on for all four alleys. Animals earned "performance scores" representing the number of trials to learn all four alleys to criterion. With the mean score of man serving as a standard, a "performance quotient" could be computed for each individual, and a mean *PQ* for the species. Pigs, requiring slightly more than twice as many trials to criterion as man, received a *PQ* of 46.5, and cats, requiring four times as many trials, received a *PQ* of 24.2. According to Fink, the performance of the animals could thus be expressed "as a *percentage* of the learning ability of man" (p. 50) !

How adequate is this comparison of the learning ability of the different species? Only a few of the more obvious flaws will be mentioned here.

(*1*) The maze was 4 ft. square, with alleys about 8 in. wide, two of them about 30 in. long and two about 60 in. long. The same maze was used for all animals, from the 2.5-in. turtle, for which the maze was originally designed, to the springer spaniel. Man, apparently because of his size, was tested with a finger-maze model of the original.

(*2*) The preëxperimental environment of the different species was not at all comparable. For example, the cats all were house-pets, the rats were laboratory bred, and the turtles all were captured in the wild.

(*3*) No attempt was made to control the massing or distribution of trials in the work here reported. For man, pig, and turtle, the number of days of testing were in the approximate ratio of 1:6:150, while trials to criterion were in the ratio 1:2:8.

(*4*) Fink's conclusions are based in most instances on means derived from ten or fewer animals, and he fails to report the characteristics of the distributions. He does mention the *PQ*-ranges for two species of tortoises (10-19, 5-33), for cats (18-34), and for humans (50-336). This hint of the variability, the small *N*s involved, and the fact that many of the species-differences are of the order of 3-8 *PQ*-points, sugggest that the differences may not be real differences.

(*5*) It is difficult to discover how Fink computed some of the individual *PQ*s he reports. The mean trials to criterion for humans was 14.8. Four trials to criterion would yield a PQ of 370, five trials 296, and so forth. How then could one of the *S*s have achieved a PQ of 336? This and several comparable errors have no place in a scholarly work.

(*6*) Rats tended to run along the outside walls, cats mewed loudly, and chicks chased flies, observations which suggest qualitative differences in the motivation of the animals. Quantitatively, Fink's claim that motivation was approximately equal in all animals, inasmuch as "the quantity of food was kept at a minimum to produce a considerable degree of hunger" (p. 23), is questionable, and he has no basis at all for assuming quantitative equivalence of human motivation (success) to that of the lower animals (food deprivation).

(*6*) The lack of equivalence of sensory equipment in the different species is

recognized: "various animal forms markedly differ in their ability to make use of different sensory clues" (p. 28). If Fink realizes that performance in a maze may depend heavily on sensory factors, then why is there no suggestion that "learning ability" depends on the particular situation employed?

The book is admittedly exploratory, and represents effort in a needed direction. Some will find the section on qualitative results intriguing, though undeveloped; others will find the attempt to relate data to metabolic factors stimulating, but naïve. It is regrettable that a work of such inadequacy should appear at a time when American animal psychology is being criticized as noncomparative. It will not help our reputation.

The State College of Washington F. D. KLOPFER

Human Development. By JOHN P. ZUBEK and PATRICIA ANNE SOLBERG. New York, McGraw-Hill Book Company, 1954. Pp. vii, 476.

This book combines in its exposition two sets of data usually kept separate. It deals not only with the growth, both physiological and psychological, of the young child into the adult, but also with later changes due to old age, or senescence. This is undoubtedly a useful combination and one which makes sense both logically and psychologically; it is surprising that until now the term 'development' has nearly always been used in a narrower sense.

There are a number of features which are characteristic of this book. In the first place, it treats the physiological side more fully than usual, dealing with such concepts as the genetic foundations of behavior, neural development, glandular development, and the like. In the second place, an attempt is made in each chapter to link up phylogenetic and ontogenetic development. This again is a somewhat novel idea which recommends itself to the reviewer. In the third place, the various areas of development (motor, sensory, intellectual, emotional, social, and so forth) are taken up in turn. In this respect, of course, the authors follow the usual pattern without attempting any major reorganization of the field.

The reviewer found some difficulty in constructing for himself a picture of the type of student for whom the book was intended. Concepts such as *IQ* are explained at great length, but concepts like factor analysis are used without explanation. Descriptions are given of the structure of the eye, but many of the early chapters assume considerable physiological background for anyone trying to follow them in detail. The student is told what a questionary, a rating scale, and a Rorschach test are, although this information surely should be taken for granted in a rather specialized book like the present one.

The coverage of different topics is distinctly uneven. The authors deal with the constancy of the *IQ* and long-term predictions of *IQ*—topics which are central to the question of intellectual development—without mentioning the Dearborn and Rothney studies, Thorndike's well known summary, or the Minnesota studies, culminating in Anderson's theory of increments. The reviewer would gladly give up the detailed description of the eye for a discussion of these much more fundamental points. Similarly, the growth and decline of intelligence is treated without reference to Cattell's concepts of fluid and crystallized ability, which seem much more important to the plan of the book than a description of the Rorschach, or a large picture of the tongue showing in detail circumvallate, fungiform, filiform, and folliform papillae. On some points, the authors do not seem to be particularly

up to date. As an illustration of conditioning, for instance, they use pupillary constriction, in spite of much recent evidence that it is an artifact. The original studies showing positive effects of glutamic acid on mental deficiency are quoted, but not later ones conclusively contradicting these findings. One study on the development of form and color perception is cited, but the very large German literature is completely disregarded. These are but a few examples of inadequate coverage. They seem to the reviewer rather serious because this is not a general textbook, but one dealing specifically with human development. It is here, therefore, if anywhere, that one must look for a discussion, or at least mention, of such important contributions as those mentioned.

Maudsley Hospital H. J. EYSENCK

Psychical Research Today. By D. J. WEST. London, Duckworth & Company, 1954. Pp. 144.

This book will be criticized both by firm believers in psychical phenomena and by skeptics. The author, a medical man, seems well versed in such psychological problems as screen memory, perception (leveling and sharpening), motivation, selective sampling, testimony, minimal cues, suggestion, automatic writing, and hypnosis. Psychical phenomena, according to West, are revealed by spontaneous cases, which he considers unverified, mediumistic cases, which are partly verified, and experimental tests, which he believes to be "well-nigh incontrovertible."

Little need be said concerning spontaneous cases. The data are subjective, spotty, and unsatisfactory. With regard to mediumistic cases, it can only be regretted that West did not heed his own warning concerning belief in instances which occurred over half a century ago. In that era, corroboration alone was thought to insure validity, no consideration was given to questions of who experienced the psychical phenomena and why, and experimental conditions, when existent, were lax. (It is of some interest to note that West is impressed by a medium whom Houdini described as a "hypocrite of the deepest dye.")

It is on experimental tests of *ESP* (telepathy, clairvoyance, and psychokinesis) that West places most emphasis, and here certain comments should be made: (*1*) West does not discuss the possibility that the concepts of 'Midas touch' (having *ESP* for a short time) and 'warming up' (preliminary trials which allow *S* to "get used to" the testing situation) may serve merely to justify the selective sampling needed for obtaining positive results. (*2*) At times the reader is inclined to believe that for West "fabulous anti-chance probabilities" supplant rather than supplement reproducibility. (*3*) Since West himself admits that in every reported *ESP* experiment there has been a "faint" possibility of an alternative interpretation, his consistent tendency to ignore the principle of parsimony is disturbing. (*4*) Displacement in scoring (being influenced by targets behind or ahead) raises objections. What are the chance expectancies under such circumstances? If such a scoring procedure were extended, and it logically might be, why should any errors occur in *ESP*? (*5*) If, as West contends, significant negative scoring (*PSE*?) requires as much knowledge of target-sequence as positive scoring, the following difficulty is encountered: Non-believers score negatively while believers score positively, and the attitude of *S* is said to be important in the obtaining of positive or negative results. It follows, then, that non-believers only think they are non-believers—they are really believers (shades of psychoanalysis)! (*6*) The logic of the conclusion

that scientists oppose *ESP* merely because it is an innovation is open to question. Although such an interpretation may hold in certain cases, opposition may also be based on the belief that the present data do not allow of an affirmative conclusion.

The history of psychical phenomena has shown advance from the supernatural to the preternatural to the natural. It is entirely possible that there are natural reasons for the "slight" *ESP*-effects found in the general population as well as the remarkable results of star-performers. Interestingly enough, it is West's own keen analysis that often leads to disbelief in cases purporting to show psychical phenomena. The author's rather sudden professions of belief in *ESP* tend to be non-sequiturs, statements based on faith rather than on fact. West himself has never obtained positive results in psychical experiments. Is he, as he claims, a poor *S?* Or is he rather a good *E?*

The State College of Washington F. L. Marcuse

Laboratory Instrumentation in Psychology. By William W. Grings. Palo Alto, The National Press, 1954. Pp. vi, 282.

Dr. Grings has given us a good textbook for an introductory course in psychological instrumentation. As far as this reviewer can determine, it is the first such text to appear. The general purpose is to introduce the student to a wide variety of instruments used in the psychological laboratory. The approach is primarily descriptive, with only moderate attention given to principles of operation. Certain areas (*e.g.* animal work) are intentionally omitted on the grounds that standard texts are available in those fields. The material is presented at the level of the student who has had only rudimentary training in physics and electronics.

The various chapters deal with Behavior Recording Systems, Timing and Counting, Audition, Vision, Other Senses, Human Learning and Perception, and Bioelectricity. Of these, the first two and the last are the strongest, and that on audition the weakest. In the chapter on recording equipment, the nature and use of kymographs, oscillographs, and other such basic components of recording systems are discussed, with attention paid to such important variables as frequency-response, reliability, and resonance. In the following chapter, a good account is given of the methods available for counting and timing of stimuli and responses. The last chapter contains a very satisfying account of the basic methods of electrophysiological recording. The section on the "General Properties of Amplifiers for Biological Recording" is particularly valuable. The chapter on audition is marked by the exclusion of certain important instruments and techniques. There is no discussion of the distinction between measurements of sound-fields in free space as opposed to sound-pressure at the ear drum. Little attention is given to the characteristics of amplifiers and other stimulating equipment, and there is no mention of certain important pieces of equipment (for example, noise-generators). In contrast to these exclusions, there is a long discussion of sound-recording and reproducing systems including, for no apparent reason, a diagram and discussion of the correct cutting angle for a disk recording stylus.

While the material is treated on a very elementary level, the book makes reference to advanced sources in the various areas. In addition, attention is focused on several of the periodicals in which new apparatus ideas may be found. Another valuable feature is the inclusion of the names of many companies producing good laboratory equipment. All of this material should be of great value to neophyte researchers.

This reviewer would have preferred to see more attention paid to general principles of instrumentation. A chapter, preceding those devoted to the special areas, in which such concepts as frequency-response, resonance, back-lash, reproducibility, and so forth, were discussed, would have been welcome. While methods of handling data are not precisely within the domain of instrumentation, a discussion of automatic stimulus-programming and data-recording systems might well have been included, together with a discussion of the related problem of reducing to a reasonable minimum the loss of data in transcription.

This book fills an important gap on the student's bookshelf. While advanced students and professional workers should find it of little value, it is a good text upon which to build an elementary course in instrumentation.

Army Medical Research Laboratory — JOHN KRAUSKOPF

Deprived Children. By HILDA LEWIS. London, Oxford University Press, 1954. Pp. 163.

The 1948 Children Act of Great Britain called for each county to establish a reception center where, on the basis of several weeks of observation and examination, physical and psychological, recommendations could be made concerning the placement of abandoned, neglected, maladjusted, and delinquent children. In this volume the author, who was the psychiatrist at the first of these centers, presents the data she and her colleagues collected concerning the background, families, behavior, and later development of the 500 children who were admitted to the center for the three years following its opening.

This book will not be found on the shelves of personal libraries; it is a reference volume and contains more statistics (71 tables) than case histories (18 pages). It provides a clear picture of the unfortunate backgrounds and families from which these children came. Although the general descriptions of behavior are useful and of interest, the children are classified according to several awkward schemes. For example, the separation of the children into the categories of *normal, normal with neurotic symptoms, neurotic, psychopathic personality, psychotic,* and *delinquent* must certainly have been accomplished with low reliability. A chapter on the influence of family and environment on the children's behavior lends support to recent discussions of the favorable effects of parental affection and the negative effects of institutional life on the mental health of children. Several startling relationships also appear in this chapter. Highly significant positive relationships were found between good mental health and dirty homes, dull mothers, and neglectful mothers. The interpretation of these findings is difficult, and little elucidation is given by the author. A follow-up study of 240 of the children two years after they left the center revealed that the proportion of children in good psychological and social condition had more than doubled, and that the children who had been placed as the staff recommended fared better than those who had not. Finally, Dr. Lewis presents her suggestions concerning the future operation of the centers.

The University of Texas — HAROLD W. STEVENSON

Introduction to Opinion and Attitude Measurement. By H. H. REMMERS. New York, Harper and Brothers, 1954. Pp. 437.

In one way or another, this college textbook deals in an introductory fashion with most of the fundamental approaches to and uses of attitude measurement, includ-

ing a variety of findings and conclusions from attitude studies. It is divided into two parts: Techniques, and Applications of Opinion and Attitude Measurement. Part I takes up in successive chapters, an introductory overview, sampling, statistics, scaling, the single question, the summated questionnaire, and "less direct measures of attitudes." Part II considers applications in business, government, industry, community interrelations, and education. Each chapter includes a list of discussion questions on the topics explicitly or implicitly covered in the chapter, as well as a list of references.

There is plenty of room left for the instructor—on several counts. Very few references date after 1948; for example, *The American Soldier* volumes (published in 1949-50) are noted only in a sentence as "a landmark in the development of social psychology" (p. 302 and footnote on p. 133). Many of the studies described are summarized so succinctly that their purpose is unclear: "Is this work included to illustrate a method or to present findings?" was a question which arose frequently during reading. The reviewer would have appreciated more examples of actual materials used in studies; the instructor will have to collect them himself. In describing methods, Remmers gives about 6 pages to Thurstone, 3 pages to Likert, 32 pages to Guttman and scale analysis (an excellent summary), and about 4 pages on the scale discrimination technique of Edwards and Kilpatrick. The approaches of Lazarsfeld and Coombs are not discussed. Another point for the instructor to consider is that the student must have a basic grounding in psychological statistics to understand a good deal of the material presented.

Perhaps what is missed most in the book is some treatment of basic attitude-measurement theory. Statistical testing of hypotheses is briefly described, but there is practically nothing on the purpose of measurement, on what is being measured, on the framing of hypotheses, on the development of specific questions used to obtain responses suitable for testing, and so forth.

It must be understood that the book contains a very large amount of information, readably presented and accurately discussed. Many instructors will find it quite suited to an introductory course in attitude measurement, for which (to the reviewer's knowledge) there is nothing comparable on the market.

University of California — RALPH R. CANTER

Manual of Child Psychology. Edited by LEONARD CARMICHAEL. Second edition. New York, John Wiley and Sons, 1954. Pp. x, 1295.

The *Manual of Child Psychology* edited by Leonard Carmichael that appeared in 1946 was found to be an extension of the earlier *Handbook of Child Psychology* that Carl Murchisom had first edited in 1931; the panel of contributors remained largely the same and to some extent the nature of the contributions (this JOURNAL, 60, 1947, 302-307). The current revision of the *Manual* is the work of almost the same panel as the 1946 publication, with the substitution of John E. Horrocks for Wayne Dennis on adolescence, a chapter by Clemens E. Benda on psychopathology of childhood to replace, perhaps, the earlier discussion of the feeble-minded child by Doll, and the addition of a chapter on social development by Harold H. and Gladys L. Anderson; the chapter by the late Kurt Lewin is extended by Sibylle Escalona to include "An Addendum—The Influence of Topological and Vector Psychology Upon Current Research in Child Development." Some of the chapters, such as Arthur T. Jersild's chapter on emotional development, have been

extensively revised and brought up to date. Lewin's chapter, written for the 1946 publication, is reproduced without change.

In terms of proportional emphasis, and as calculated from the numbers of pages, about a fourth of the material surveys *growth periods* (Carmichael on the fetal period, Ruth M. Cruikshank, Karl C. Pratt, and Arnold Gesell on infancy, and Horrocks on adolescence); another fourth is given to *individual differences* (Florence L. Goodenough on measurement of mental growth, Harold E. Jones on environment and mental development, Lewis M. Terman and Leona E. Tyler on sex differences, Catherine Cox Miles on gifted children, and Benda on psychopathology. More than a third of the book presents *types of development;* physical growth (Helen Thompson), learning in children (Norman L. Munn), language development (Dorothea McCarthy), character development (Vernon Jones), emotional development (Jersild), and social development (Anderson and Anderson). The remaining 13% is devoted to discussions of *methodology* (John Anderson, Margaret Mead) and *theory* (Lewin and Escalona).

The format of the second edition is very similar to that of its predecessor, and it is apparent to this reviewer that the editor and contributors have preserved the scholarly and scientific integrity of the *Manual* while bringing the survey up to date. The book should be useful not only as an ultimate reference in child psychology, but as a carefully developed guide to further research. Its retail cost is twelve dollars which may suggest that its suitability as a text required of students should be carefully considered.

School of Medicine
Louisiana State University

T. W. RICHARDS

Nebraska Symposium on Motivation: 1954. Edited by MARSHALL R. JONES. Lincoln, University of Nebraska Press, 1954. Pp. x, 322.

Here is another published record of a symposium—this one on motivation, sponsored by the Department of Psychology of the University of Nebraska, and financed by a training grant from the United States Public Health Service. The competent contributions of Atkinson (on imaginative productions), Farber (on anxiety), Festinger (on social behavior), Klein (on cognitive control), Nissen (on animal studies), and Ritchie (on certain formal problems) are diverse both in aspect and in scope. One wonders what function such a volume is expected to serve.

M.E.B.

Thinking and Speaking: A Symposium. Edited by G. RÉVÉSZ. Amsterdam, North Holland Publishing Company, 1954. Pp. 206.

Révész has brought together ten essays on the relations between talking and thinking which appeared originally in *Acta Psychologica* (1954). The ten contributors write in three different languages, represent five different disciplines, come from eight different countries, and illustrate ten different ways of phrasing the question. Since the integration of these disparate views is left to the reader (and there is not even an index to help him), it is difficult for a reviewer to summarize what progress the book represents toward the solution of this old, familiar problem. The flavor of the book is best preserved by a consideration of each in turn.

First, the psychologists, which include Révész himself, Piaget, Kainz, Eliasberg,

and Cohen. Révész's contribution, "Denken und Sprechen," comes first, is longest, and makes the most ambitious attempt to resolve the dilemma, the two horns of which are the identity theory and the dualistic theory. Révész rejects both. Thinking and speaking cannot be identical because their ontogenesis is different, because words are often inadequate expressions of thoughts or emotions, and because eloquence is not perfectly correlated with intelligence. Thinking and speaking cannot be two distant and independent processes because there are many congruities between them and because disturbances of speech and disturbances of thought so often occur together. The way out of the dilemma is called *Dualistische Einheitslehre,* a dualistic theory of unity, which says that thinking and speaking are two distinct processes that are completely dependent upon one another. The existence of one always presupposes the existence of the other. The bulk of the article is an elaboration of this theory.

The task for Révész, therefore, is to show that speaking and thinking obey different laws although neither occurs without the other. The argument for their obedience to different laws is that they serve different functions and have different structures. Speaking serves the function of communicating, while thinking serves the function of cognizing or understanding. The structure of speaking is given by grammar and syntax, while the structure of wordless thought is unknown. The argument that one never occurs without the other falls into two parts. First, the contention that speech cannot occur without simultaneous thought is supported by references to sensory aphasia, to the congenitally deaf, and to backward children. Second, the contention that thought cannot occur without an ability to speak is supported by references to the effects of progressive deafness, the subconscious pattern that language imposes even upon nonverbal thought, and the absence of evidence to the contrary. If, however, one feels that Révész's original definitions of thinking and speaking are unsatisfactory, the elaborate argument that follows is apt to seem beside the point.

The article by Révész serves as an introduction to the others. Piaget concludes that "Speech and thought are . . . supported by each other in a perpetual interaction, but either depends on intelligence, which is anterior to and independent of speech." Kainz agrees with Révész in general and points to the involvement of related parts of the brain in both activities; he is concerned also with the effects of naming percepts and the consequences of such categorization for supposedly nonverbal processes. Eliasberg seems to disagree somewhat with Révész and to argue that aphasia simply reduces the accessibility of speech-cues without affecting the ability to form creative abstractions. Cohen reviews some of the literature on this problem and reports an experiment that seems to show that verbalization is better if a delay is inserted between the assignment and the performance of a writing-task.

A philosopher's approach is offered by Jørgensen who notes the existence of problems that cannot be solved by mere manipulation but demand imagining or intending changes in the problem-situation. Such imagining presupposes the use of symbols, and symbols presumably play a great part in human thinking, but Jørgensen is doubtful that linguistic symbols are always necessary. Buyssens approaches thinking and speaking from the linguistic standpoint. He explains with care that the interpretation of linguistic categories does not require us to assume a similar grammar within our thoughts. Van der Waerden, the mathematician, reports his own introspections about mathematical thinking. He holds what Révész

would call the dualistic theory and he finds that language has little influence or importance.

Goldstein's paper adds little to his earlier work but does serve as a convenient summary of his point of view. Goldstein is less interested in the relation of thinking and speaking than he is in the distinction between normal and pathological accomplishment in either area. The distinction is found in the well-known dichotomy of concrete and abstract behavior; it matters little to him how speech and thought relate as long as both are seen to reflect one general, underlying, organismic activity. In this same vein of psychiatric neurology, Gruhle is much more cautious and more critical in his interpretations of the behavior of aphasics. Gruhle reconsiders the material collected by Gelb and Goldstein in support of their theory of color-naming amnesia, and he argues that the data do not show irrefutably that the naming-difficulty actually is accompanied by a disturbance of the thought-processes instrumental for categorizing. Depending upon the clinical data one selects and the definition one adopts of thinking and speaking, one can either affirm or deny the possibility of speaking without thinking or of thinking without speaking.

A summary evaluation of such collection is not easy to make. Most readers will probably develop a kind of semantic restlessness as they go through it; speaking and thinking are so vaguely and variously defined that the task is like relating the smoke to the fog in a cloud of smog. Nevertheless, the authors labor mightily to clarify the issues and it is not obvious what should be done to improve on their efforts. It is possible, of course, that they have posed an impossible question, which might be suspected from the very fact that it is such an ancient and honorable one. Suppose, to shift the issues into a different sphere, that a group of physiologists wanted to discover the relation between walking and swimming. They might note that the two are not identical because they serve different functions and they might add that the two are not completely distinct because a man who has lost his legs cannot do either. They might conclude that walking and swimming were distinct but highly dependent functions. We may even suppose that the physiologists resolved the question to everyone's satisfaction. Even so, it is difficult to see what the answer would contribute to their understanding of the basic physiology involved in either process. If the answer to a question is not enlightening, then the question must be a poor one. Many people, however, particularly clinical and anthropological people, persist in asking what relation speaking bears to thinking, and it is scarcely a service to tell them they should not want to know. For such people, therefore, this book will be a welcome source of ideas and opinions.

Harvard University

George A. Miller
Eric H. Lenneberg

Social Learning and Clinical Psychology. By Julian B. Rotter. New York, Prentice-Hall, 1954. Pp. xvi, 466.

Dr. Rotter's interesting volume contains more than a new theory of learning. It also contains a highly perceptive analysis of the relationship between psychological and physiological constructs, a description and critique of existing psychological tests including those of the projective type, a very practical survey of psychotherapeutic methods, and much good advice on the management of children. Since, how-

ever, the reviewer's primary interest lies in the field of learning, his comments will be confined largely to that aspect of the book.

The author's main concern is with the clinical field, and his basic purpose is "to arrive at a systematic theory from which may be drawn specific principles for actual clinical practice" (p. viii). Because of the conflicting pressures under which clinical psychologists labor—the pressure to be 'scientific' and the pressure to achieve practical results with patients—their work and thinking often are unsystematic and confused. Among their faults are a disregard for basic theory and experimentation, overgeneralization, the use of tests of dubious validity, reliance on the nosological approach, failure to define their terms (*e.g.* frustration, anxiety, catharsis, schizophrenia) with precision or with definite referents, and a consequent inability to communicate unambiguously with each other. In the main, Rotter's criticisms appear to be soundly based, adequately documented, and worth the attention of every clinical psychologist who is willing to face their implications. It is evident, however, that such faults are not limited to the clinical division of psychology.

Early in his book, Rotter discusses the criteria for a language of description in clinical psychology. Reliable terms are those which have objective behavioral referents. The ideal scientific language has "as many terms as are necessary for reaching any given level of description and no more" (p. 48), a statement that helps to account for the small number of variables which Rotter himself employs. The author strongly prefers operational definitions, and insists that the difficulty of defining operationally the concepts of clinical psychology is no excuse for "remaining in a state of semantic confusion, controversy, and fuzzy thinking" (p. 51). He then discusses the language of faculty psychology, of personality types, of personality traits, of psychoanalysis, and finally of learning theory.

It is the language of learning theory which Rotter selects as most useful for clinicians, but he finds no existing theory satisfactory for his purpose, which is the conceptualization and prediction of *human* behavior. Hullian *S-R* theory is basically a "theory of learning of lower animals" which "is not at present adequate for predictive purposes in the field of personality study" (p. 83). Its major fault seems to lie in the apparent impossibility of measuring drive and drive-reduction in human *S*s. Expectancy or cognitive theories afford little bases for prediction—how, for example, does one determine beforehand the strength of an expectancy?—and they do not give a sufficiently important rôle to reinforcement.

Rotter's own theory represents an integration of the expectancy and reinforcement points of view, but with qualities not to be found in either approach as it exists at present. The three basic concepts of the theory are: (*1*) *behavior potential*, "the potentiality of any behavior's occurring in any given situation or situations as calculated in relation to any single reinforcement or set of reinforcements" (p. 105); (*2*) *expectancy*, "the probability held by the individual that a particular reinforcement will occur as a function of a specific behavior on his part in a specific situation or situations," which "is independent of the value or importance of the reinforcement" (p. 107); (*3*) *reinforcement value*, "the degree of preference for any reinforcement to occur if the possibilities of their occurring were all equal" (p. 107). The relationships among these variables are set forth in "basic formulas," such as $B.P._{xs_1} R_a = f\ (E_x\ R_a{}'\ {}_{s_1}\ \&\ R.V._a)$ which reads: the potential for the occurrence of behavior *x* in situation *1* in relation to reinforcement *a* is a function

of the expectancy of the occurrence of reinforcement *a* following behavior *x* in situation *1* and the value of reinforcement *a*. The sign & is used between *E* and *R.V.* because insufficient data preclude a more precise formulation, but "it seems fairly certain . . . that this relationship is a multiplicative one" (p. 108). For prediction, another variable also is important, the "psychological situation," by which is meant "the manner in which a person perceives a given situation" (p. 200), but this variable appears not to enter into the formulas. These formulas eventually are expanded to include the factor of generalization, a concept of great importance to Rotter. Thus, the expectancy of reinforcement in a given situation depends in part on experiences of similar reinforcements in that situation, and of the same or similar reinforcements in similar situations.

Expectancy of reinforcement is Rotter's closest equivalent to a habit construct. He follows "the Hullian tradition in attempting to predict [its] strength . . . on the basis of decreasing increments and in specifying variables that affect [it] such as frequency of occurrence, order of occurrence, and generalization" (p. 80). Expectancy is readily capable of measurement in absolute as well as in relative terms, *e.g.* in experiments of the familiar level-of-aspiration type. Reinforcement value is a somewhat unusual concept, though it seems affiliated with Lewin's notion of valence and Tolman's idea of demand. It signifies "a mathematical statement of a preference on the part of the subject . . . for this reinforcement or objectively describable event to take place" (p. 149). It can best be measured in relative terms; for example, by asking *S* to choose between one dollar and ten dollars, or to say whether he likes this toy better than that. As the formula for behavior potential indicates, expectancy and reinforcement value share in determining behavior. Some of their relationships, as Rotter describes them, have considerable interest. For example, a child may become delinquent, not so much because the reinforcements attendant upon delinquency have greater value for him than other reinforcements have, but because he has relatively little expectancy that other types of behavior will be positively reinforced. In other cases, delinquent behavior may occur because of the high value of its associated reinforcements, even though the expectancy of obtaining them is not great.

The noteworthy omission, if it may be so termed, in Rotter's theory is the absence of any independent variable labeled drive, need, motive, or the like. A principal reason for this omission is Rotter's complete rejection of the idea that human needs, such as dependence, recognition, and dominance, have any but a possible historical relationship to physiological drives. Such drives may be capable of measurement and quantification, but human needs are not. Hence they cannot usefully enter into the formulation of an operational theory. Another reason for the omission of the drive variable is Rotter's adherence to a strictly empirical law of effect, in pursuance of which he defines reinforcements as "identifiable events that have the effect of increasing or decreasing the potentiality of some behavior's occurring" (p. 148). Anything that we would term a reinforcement implies both a need and a goal, but these are inferred solely from the behaviors observed and their modification. It is sufficient if we regard behavior as directed toward certain reinforcements, thereby defined as positive, and directed away from other reinforcements, thereby defined as negative. The probability of the occurrence of any behavior is a function of the individual's expectancy of reinforcement and the

value of the reinforcement to him. Prediction will not be improved by introducing the concept of need or drive.

This approach is a somewhat novel one, and at first glance may seem to do unnecessary violence to traditional ways of thinking, especially about clinical problems. Certainly many psychologists talk as though they believe that the term 're-inforcement' is virtually meaningless without the assumption of a previously existent need which is satisfied or reduced. It is entirely possible, however, that the construct of need, drive, or motive is dispensable, at least for predictive purposes and from a strictly operational point of view. Indeed, psychologists like Sheffield and Harlow, who cannot accept the restriction of reinforcement to drive-reduction, might find that Rotter's approach suggests new ways of relieving them from that incubus.

Rotter discusses his basic concepts in detail, describing many experiments illustrative of the kinds of experimental design which his theory suggests, and the ways in which his variables can be measured. Almost all of these studies are doctoral dissertations at Ohio State University and unfortunately are not published elsewhere. They at least effectively demonstrate that the theory is capable of generating propositions which can be put to actual experimental test.

Criticism of Rotter's theory is easy, as indeed most criticism is. One can note his exclusive concern with the behavior of the older child and the human adult, and point out the inapplicability of his concepts to infant and animal behavior. One can stress the fact that thus far his propositions provide only a fragment of a comprehensive learning theory, only a 'miniature' system at best. For example, Rotter does not deal anywhere with such topics as conditioning, spaced versus massed learning, retention, retroactive inhibition, and so forth. It might even be said that we do not have here a theory of *learning* at all, but only a partial description of selected aspects of human behavior, which is presented in terms of a few variables, mainly derived from Tolman, and—in its way unique—is as much an oversimplification as any zoömorphic theory of human activities could be. Yet it seems to the reviewer that so far as possible one had best regard a theory in accordance with the avowed intentions of its author. Rotter, as we have seen, aims only to develop a few learning concepts, a little of a *language* of learning theory, which should be useful to clinicians in their efforts to conceptualize the behavior of their patients in a systematic and consistent way. He presumably hopes that at least some clinicians, after they have become familiar with it, will prefer it to other languages, including those of psychoanalysis, of most current personality theory, and of Hull, Skinner, Guthrie, and Tolman as well.

The reviewer often has wondered why some psychologist did not integrate within a single learning theory the concepts of expectancy and reinforcement. It seemed to him that their customary separation was due principally to accidental factors, and in no way reflected any basic incompatibility between them. It also seemed to him that clinicians in particular would welcome such a union, provided expectancy could be divorced from what they frequently regard as its over-intellectualized connotations, and reinforcement freed from its bondage to *S-R* postulates. It is just this integration that Rotter has tried to achieve. From this point of view, if from no other, his social learning theory merits a sympathetic inspection by clinicians and learning theorists alike.

New York University LELAND W. CRAFTS

Current Trends in Psychology and the Behavioral Sciences. By JOHN T. WILSON and Others. Pittsburgh, University of Pittsburgh Press, 1955. Pp. xiv, 142.

Here is a heady topic, and the participants in this eighth annual conference grapple with it bravely. J. T. Wilson calls attention to a "sincere and effective *rapprochement*" between psychology and related disciplines which must lead to a "striking synthesis of knowledge in the behavioral science domain" (p. 23), although 51 references, from Allport to Zener, somehow fail to carry the point. C. S. Ford foresees the disappearance of barriers between psychology and anthropology, predicting the development of a "truly general science of human behavior and social living" (p. 37). B. F. Skinner announces his readiness to apply promising recent advances in the field of learning to the betterment of our educational system—could less have been expected of one who has taught pigeons to play ping-pong? After all, "the species of organism" has been found to make "surprisingly little difference" (pp. 43-44). What *is* important, apparently, is reinforcement—frequent and immediate—and Skinner has a box that will supply it. G. Bergmann re-ponders some familiar problems of reductionism. F. A. Beach discovers that development is influenced both by hereditary and by environmental determinants (with interdisciplinary examples). Finally, K. H. Pribram describes the effects of certain brain-lesions on the behavior of the monkey in support of the proposition that "neuropsychological" experiments will provide a "common framework relating physiological and behavioral science" (p. 140). The speakers are introduced by R. A. Patton, who chooses the occasion to propose a definition of "psychological stress."

Dublin, Ireland

T. S. KILLARNEY

The Psychology of Politics. By H. J. EYSENCK. London, Routledge and Kegan Paul, 1954. P. xvi, 317.

This is a book which leaves the reviewer gasping for adjectives. Parts of it are models of lucidity, while others are so incredible that one wonders whether some practical joke has been attempted. In essence, the book brings together two previously distinct lines of research by the author, one on dimensions of personality (in which two factors—neuroticism and introversion—have been isolated), and a second on primary social attitudes (reduced by factor analysis to two orthogonal dimensions—radicalism-conservatism and "tough-mindedness"-"tender-mindedness"). Eysenck maintains that " 'tough-mindedness' is a projection on to the field of social attitudes of the *extraverted* personality type, while 'tender-mindedness' is the projection of the *introverted* personality type" (p. 174). It seems fair to view this attempt to link personality variables with political ideology as a tough-minded (without quotes) attack on the problems previously dealt with in *The Authoritarian Personality.*

The first three chapters of the book consider certain relationships between social class and attitude, problems of polling, and attitude-measurement. The presentation is sound and lucid—elementary in the best sense of the word. Budding social psychologists might be advised to study these chapters carefully; the author's ability to summarize pithily and without distortion a highly varied literature has seldom shown to better advantage. In the next four chapters (the meat of the book) the work of Eysenck and his students provides the warp, and the highly selected woof

of other research is woven into a uniquely Eysenckian pattern. It may safely be predicted that the extent of the reader's acceptance of what appears on the surface to be a highly plausible argument will be inversely proportional to his familiarity with the material on which it is based. These pages are grounded upon samples which are at best atypical and at worst highly aberrant, upon the use of scales which do not measure what they are purported to measure, and upon a misleading analysis of results. It is to be regretted that Eysenck does not apply to his own work and to that of his students the same critical standards which he applies to the work of others.

The eighth chapter ("A Theory of Political Action") represents an attempt to relate the material of the preceding four to Eysenck's conception of Hullian learning theory. Eysenck contends that "the concept of *attitude* corresponds in every detail to the concept of *habit* as discussed by Hull" (p. 246). Despite the title of this chapter, it contains no reference to the considerable body of literature on political behavior. A thorough knowledge of that literature might have led the author to a greater awareness of the complexity of social life and made him less willing to assert that an analysis of "political behavior in terms of ${}_sE_r$ as a multiplicate function of ${}_sH_r$ and D appears sufficient to account for our Radical and Conservative group of attitudes" (p. 259).

The book displays a good deal of carelessness, not only in conceptualization, but in such basic matters as the reporting of data. One of the simplest tests (which sometimes yields astounding results) is to make a 'reviewer's run test' and recheck the addition in tables. A limited sample consisting of the first six tables was examined. Four contain obvious errors in addition (pp. 18, 19, 20, 21). These flaws are not crucial to Eysenck's argument, but they lend little confidence in the care which he devotes to statistical analysis (although some of my more sophisticated colleagues maintain that a knowledge of addition is not really essential in factor analysis). The sloppiness of the reference-system also is annoying. A bibliography of some 500 titles is given, but most of these are not referred to in the text, and many references are made in the text to works which are not cited in the bibliography (*e.g.* to Guilford, p. 176, and to Hildebrand, p. 177). Unless one knows the material referred to extremely well, it is difficult to trace direct quotations. For example, Barron is directly quoted on p. 184 but no indication is given as to which of the three references to Barron in the bibliography is the source. There are numerous references to unpublished theses written by Eysenck's students but they are hard to track down. For example, there is a reference to "Melvin (1954)" on p. 276, but Melvin is listed in the bibliography only as the author of a London University Ph.D. thesis of 1953, and, according to a letter received from the library of that university, there is no such thesis on file.

A writer is privileged to cite whatever references he may choose, but one wonders at some conspicuous gaps. Despite the importance given to introversion and extraversion, there is no mention made of the pioneer study of F. H. Allport and D. A. Hartman (*Amer. Pol. Sci. Rev.*, 19, 1925, 735-760) on the possible relationship of these personality traits to tough-mindedness and tender-mindedness. It might be noted that the results of this suggest a relationship between radicalism and tender-mindedness, while Eysenck holds a different view. More direct evidence of poor taste is exemplified by failure to acknowledge credit where demanded by ordinary

standards. For example, on p. 222 a rigidity scale is reproduced with no indication of its source. All 22 items are identical, word for word, with those of an unpublished scale prepared by Gough and Sanford in 1952.

Eysenck's lucidity may be traced in part to the fact that he models himself upon good writers. Compare, for example, the following excerpts: "The scapegoat . . . theory . . . maintains that the individual high in prejudice has a certain amount of hostility or aggression that he has not been successful in reducing or acting out against the original object of aggression. The theory holds further that the person succeeds in reducing his hostility by displacing or redirecting it upon the more or less helpless members of minority groups in the form of prejudiced behavior" (G. Lindzey, *J. Abnorm. & Soc. Psychol.*, 45, 1950, 296). "The scapegoat theory maintains that the individual high in prejudice has a certain amount of hostility or aggression which he has not been successful in expressing or acting out against the original object of aggression. . . . The theory holds further that the person succeeds in reducing his hostility by displacing or redirecting aggression against the more or less helpless members of minority groups in the form of prejudiced behavior" (Eysenck, p. 215).

The chief value of the book appears to lie in the first three chapters which would be useful in an introductory course in social psychology. The remainder might be used in courses on methodology as an example of how *not* to analyze and report data.

Research Center for Human Relations
New York University

RICHARD CHRISTIE

INDEX

By H. W. STEVENSON, University of Texas

(The names of authors of original articles are printed in CAPITALS; of authors of books reviewed, in roman; and of reviewers, in *italics*.)

SUBJECTS

(References in *italic* figures are to reviews.)